UNIVERSITY OF SUSSEX

Essays on Beliefs, Democracy and Local Labor Markets: An Empirical Examination for Peru

Edgar Salgado Chavez

Submitted in partial fulfilment for the degree of Doctor of Philosophy in Economics

University of Sussex

September 2017

Abstract

This thesis presents three empirical chapters on local labour markets, mineral booms, beliefs, conflict and uncertainty. All the analysis was conducted using Peruvian data and context.

The first chapter finds that Peruvian individuals exposed to violent events during their impressionable years trust less government institutions, and feel less identified with their neighbours, while more identified with religious groups. The estimated effect is small and heterogeneous depending on the identity of the perpetrator. The effect on identification with groups of population is also heterogeneous by the indigenous origin of the individuals. Owners of an agricultural plot embedded in a cooperative setting at the local level exhibit even smaller levels of identification with their locals while higher levels of identification with their ethnic group. In line with recent literature, these findings suggest that conflict has a small but persistent effect on the formation of trust and identity, which is a central feature to understand the interaction between culture and institutions, and ultimately to understand the persistent consequences of wars.

The second chapter studies the relationship between democratic beliefs and economic uncertainty. I explored whether uncertainty experienced during the impressionable years of the individuals is a key factor behind the formation of the democratic beliefs. Results showed that this type of uncertainty had no effect on the determination of democratic beliefs. Combining uncertainty with the exposure to authoritarian regimes did not change the result. This result is robust to different definition of rural individuals, the interaction of uncertainty and degree of experienced authoritarianism, and different formative periods. Current uncertainty, on the other hand, was unable to fully explain the formation of democratic beliefs.

The final chapter investigated the local labour effects of mining booms. Using two rounds of population census for 1043 districts in Peru I documented that large-scale mining activity had a positive effect on local employment over 14 years. The effect was differentiated by industry, skill and migration status. Employment grew by 4% faster by one standard deviation increase in the mineral prices. Both high and low skilled workers enjoyed similar employment increase, however only low skilled workers experienced a decline in unemployment. Using data from 10 annual household surveys I found that, consistent with a model of heterogeneous firms and labour, wages for low skilled workers in districts close to the mining activity was 5% higher by every standard deviation increase in the index of mineral prices. Additional evidence with the census data suggested that to a large extent locals working in the mining or the agricultural sector filled the new employment opportunities. Together these findings suggest that large-scale mining activity increases the demand for mining and agricultural local employment, and the wages in the local economy.

Contents

List of Tables

List of Figures

Acknowledgments

Paolo Masella and Andy McKay guided me through the elaboration of this document, and to them my inexpressible gratitude not only for the rigorous academic advise but also the support and encouragement. Richard Dickens was also my supervisor at an early stage and to him I also want to extend my gratitude. Sebastian Sotelo has been an exquisite friend who uncountable times has granted me advise that goes beyond valuable. The last chapter of this thesis owns too much to his comments and suggestions. The Sussex community, and in particular my fellow PhD candidates, also provided a vibrant space for earlier discussions that in many ways fueled the matter that shapes the body of this document. I obtained the data on mining production from Nemera Mamo and Sambit Bhattacharyya. Finally, I want to thank my mother, whose effort to put me in the direction that leads to a PhD clearly overweighs the effort required to obtain such academic degree.

General Introduction

Could a war determine how individuals trust their peers or institutions? Would economic uncertainty shape democratic beliefs? What are the long run local effects of mining booms? This thesis groups three chapters which aim to answer these questions using the Peruvian context between 1990's and 2000's.

Seemingly unrelated as these topics seem to be, they represent some of the key events that depict recent Peruvian history. In the period of analysis, Peruvian society has been exposed to events that carry the force for deep transformations: economic booms, sustained uncertainty and wars. The first and second chapters are linked through the study of shocks during the impressionable years of the individuals. The topic each chapter addresses, however, varies. Conflict, trust and identity are the target of the first chapter, while democratic beliefs and uncertainty define the agenda in the second chapter.

While economic booms have the potential to transform not only the aggregate economy, they also shape the local economy, and this is the focus of the thrid chapter. In the preamble, I discuss the historical context that links the three topics: local labor markets, democratic attitudes and beliefs about trust and identity, while here, I provide a succinct summary of what every chapter presents and explain their relevance for the Peruvian context.

I do find that the exposure to conflict has a statistically significant, although small, effect on the level of trusts that individuals express, and also a small effect on their identification with their peers. And I also find that democratic beliefs are not related to economic uncertainty; in particular, uncertainty during the formative years of the individuals does not seem to have any explanatory importance in the determination of how much individuals like democracy. I also find that mining booms increase employment and wages at the local level. A few illustrations highlight the importance of these topics.

Peru is a country that witnessed one of the last civil conflicts in South America, and it is fair to say that such conflict captured the attention not only of Peruvians. After serving 20 years in prison, Lori Berenson, an American citizen accused of terrorism for participating in a plot that aimed to blow up the Peruvian Parliament in 1995, returned to the U.S.[1] This

[1] http://www.nytimes.com/2015/12/03/world/americas/lori-berenson-returns-us-peru.html?_r=0

event revived a vitriolic discussion in Peruvian media about the consequences of the most recent civil war, and the punishment that should be applied to those found responsible for it. Mrs Berenson was frequently harassed on the streets of Lima, while she was waiting to finish her parole. This ultimately only reflects how much that violent period between 1980 and 2000 still shapes the political discussion in Peru. In the openning chapter I analyze its effect on the degree of trust expressed by individuals who were somehow exposed to it in their locality.

There is also the political instability that has characterized Peru in recent years. 195 years of republican history have not been enough to guarantee a stable democratic regime in the country. Pedro Pablo Kuczynski just recently won the presidency for the 2016-2021 period, in the fourth consecutive democratic election after the last authoritarian period: 1992-2000. Four consecutive democratic elections may seem little, but this has been celebrated in Peru. Mr. Kuczynski won last minute support from the left after openly confronting his adversary and accusing her of being the representation of a likely return to authoritarian times[2]. Yet, almost half of the country was willing to support the adversary, Keiko Fujimori, the daughter of the dictator who ruled the country between 1990 and 2000. In the second chapter, I try to understand why so many Peruvians would still vote for a non-democratic option.

Finally, Peru is a middle income country whose economy has depended on the exploitation of mineral resources since the colonial period. The extraction of gold, silver or copper has shaped the political discussion in the country since its independence to this day. A few months ago, Maxima Acuña, a subsistence farmer living in highlands of Cajamarca in the north of Peru, managed to stop the expansion plans of Yanacocha, the largest gold mine in Peru. She won a litigation case that prevented the mine to expropriate her land, and she refused to sell the land, as well. Many observers[3] have seem Mrs Acuña's case as the latest expression of the historical tension between the quintessential Peruvian export and the local communities. Without making any grandiose claims, in the last chapter of this thesis I aim to contribute in the understanding of this interaction that lays at the core of Peruvian economic policy.

[2]http://www.newyorker.com/news/news-desk/a-surprising-coalition-brings-a-new-leader-to-peru

[3]https://www.theguardian.com/sustainable-business/2016/apr/21/peru-farmer-wins-battle-newmont-mining-corporation

Perhaps these topics seem relevant in any context, however there is one last point to emphasize in this regard that also determines the choice of Peru as context for the analysis. Since early the 1990's, there has been a growing availability of high quality data in Peru. For instance, unlike many other Latin American countries, Peru has detailed household surveys running from 1997. It has been a constant preoccupation among Peruvian official bodies to provide detailed amounts of data. This thesis greatly benefits from this, and somehow provides a byproduct of relevance: a data set that combines some of the determining factors that shaped Peru in recent years.

Preamble: The Historical Context

Peruvian history has been permeated by the interaction between its dependency on natural resources, the civil conflict that devastated the country and a feeble democratic system. In this section I try to historically connect all these events through the description of recent historical accounts with glimpses to the far colonial past.

On the 5th of April 1992, president Alberto Fujimori shut the parliament and inaugurated the last autocratic period known in the republican history of Peru. This drastic measure has been recorded as the last cuop d'état in an long history of coups d'état in the 195 years of existence of the Peruvian republic. The difference this time was that it was self-inflicted.

The country was submerged in chaos. The GDP per capita was 6,803 Soles, the national currency, the lowest point in years, comparable to the GDP per capita in 1959, 6,591 Soles. 34 years separated 1959 and 1992, and yet the GDP per capita had grown a mere 3.2% during the whole period. On an annualized basis, this average represents a 0.09% annual increase between 1992 and 1959.

History, of course is not linear, and this was the consequence of a turbulent period in Peruvian history which delivered to the society, among other things, a devastating agrarian reform (1969), a war period (from 1980) that cost the lives of approximately 69,000 Peruvians by the time it ended, an hyperinflationary period (1987-1990) that now populates some pages of economics textbooks, the succession of military autocratic regimes (1962-1963, 1968-1980) intertwined with the most disastrous economic performance in recorded history, which belongs to president Alan Garcia (1985-1990).

It is perhaps this agitated period what led to the validation of Fujimori's autocratic regime in 1993. Following the coup, the president elaborated a new constitution published on the 4th of September 1993 and submitted to the population's approval through a referendum 56 days later, on the 31th October. That day 52% of the electorate voted in favor of the new constitution, an event that in a way symbolizes not only a new era, charged with civil authoritarianism, economic success and the consequences of war, but also the time frame for the analysis done in this thesis.

Alberto Fujimori's ascent into power in 1990 was a reaction to the widespread upheaval

in the political and economic scene. Alan Garcia's government from 1985 to 1990 is remembered for the soaring inflation, upsurge in the violence led by the terrorist group *Sendero Luminoso*, increase in unemployment and expansion of the informal economy. Fujimori won the presidency by opposing the economic plan proposed by his contender, Mario Vargas Llosa, who wanted to shock the economy by eliminating all price and exchange rate controls, among other drastic economic measures. Necessary measures fiercely opposed by Fujimori during the electoral campaign, which however had to be implemented just 11 days after being sworn in as president, on the 8th of August 1990[4]. Fujimori was re-elected president in 1995 under high rates of approval from the population who saw his autocratic style as a necessity for this troublesome period.

Alberto Fujimori's government came to and end in 2000 amid widespread corruption and soaring poverty. The president had ran for a third mandate in 2000 elections, contravening the constitution. The elections that year were marked by a sense of fraud, and the other candidate, Alejandro Toledo, withdrew himself from the runoff. The controversial election results set Fujimori for a third mandate, facing strong opposition from the public, which ultimately led to his resignation, while he was in Japan, via fax. The parliament appointed Valentin Paniagua as provisional president until new elections were held in 2001. Alejandro Toledo this time won the presidency, and from that moment the country gained some political stability and poverty started to decline, something that lasts until to day.

With the desire to vindicate the atrocious economic performance of his first presidential period, Alan Garcia run for the presidency again in 2006 and defeated Ollanta Humala, a leftist candidate who put himself as the Peruvian impersonation of the leftist political movements that populated the South American continent during early 2000's. Garcia's presidency followed many of the policies already implemented by the government of Toledo and poverty kept falling under his mandate. Ollanta Humala ran for the presidency again in 2011, this time with a less radical political agenda, and defeated the daughter of Fujimori, Keiko Fujimori, who was also running for the presidency that year.

[4]In a poignant televised speech, the finance minister of that time, Carlos Hurtado Miller, enumerated the drastic changes in prices for some of the most important goods (milk, gasoline, bread), one by one, to the shock of the population. The measures announced that day were so drastic that the minister had to finish his speech with one of the most remembered pleas in recent history: *and may God help us*.

If something characterizes the presidential regimes in Peru, it is their low rate of approval among population. The two exceptions in modern times are Alberto Fujimori and Valentin Paniagua. The first had an average approval rating around 50% until almost the end of his period, while Paniagua's approval rating reached an astonishing 72% at the beginning of his short presidency. Alejandro Toledo finished his presidency with an approval rating of 18.5%. Alan Garcia's government from 2006 to 2011 reached an approval rating of 33% despite the economic boom the country was experiencing. President Ollanta Humala's approval rating was also close to 18% during most of his period as president.

Although politically there is a general dissatisfaction among Peruvians, economically, the country thrived. All presidents from Fujimori himself followed a policy that privileged openness to markets and less participation of the government apparatus in the economy. In recent years some redistributive policies were implemented, which resulted in the reduction of the poverty rate to 21.8% in 2015. Demographically, the 1993 population census revealed that the 65% of the population was urban, while by 2007, that percentage had increased to 72%. The urban-rural gap in the poverty reduction is still evident. In 1994, 39.6% of the urban population was poor, while 71.1% of rural households were poor. By 2015, the urban poverty rate declines to 14.5%, while the rural poverty headcount was 45.18%.

From 1993, the country was an economic success in many levels. The average GDP annual growth rate from 1993 to 2012 was 5.3%, while GDP per capita grew on average 3.9% every year from 1993 to 2012. The life expectancy for a random individual in 1994 was 68 years, while in 2012 was 74. Economic prosperity has, nonetheless, also meant higher inequality. The Gini coefficient grew from 0.39 in 1997 to 0.44 in 2012. The national extreme poverty rate by 2015 was 4%, which is a remarkable outcome for the country, however, the urban-rural gap is even more evident here. Just 1% of the urban population lives under the extreme poverty line, while 13.9% of the rural population is still extremely poor in 2015. The rural highlands and jungle account for most of it. Extreme poverty in the rural highlands is still 16.5%, whereas in the rural jungle, the extreme poverty rate reaches 10.5% of the population.

Partly this prosperity has come as a consequence of the boom experienced in the prices of the raw materials the country exports: gold, silver, copper, lead and zinc. Historically, Peru-

vian economy has been dependent on the exports derived from minerals. This dependency dates back to the times of the Spanish conquest. It is famous the anecdote about the ransom paid by the Inca (the ruler of the Inca empire) Atahualpa to the Spanish conquerors at their arrival. Hoping to be released from his captivity, the Inca Atahualpa offered a room filled with gold and silver to Francisco Pizarro. This fact marked the beginning of a kind of gold fever that characterized the Spanish occupation from 1532 to 1821.

At the macroeconomic level, the mining sector is the main tax contributor. Yet, the local effects of its activities remain disputed. The abundance of natural resources in Peru has been historically surrounded by criticism. Dell (2010) for instance, finds that the exploitation of a mercury mine between 1537 and 1812 in colonial Peru still had some negative effects on the level of consumption and children height in 2001.

With the number of mining projects boomed from the 1990's, the number of social conflicts derived from this boom also boomed. As stated by Oxfam (2009) the root of the conflict has its explanation in the few benefits that local communities perceive from mining revenues. On the contrary, communities live under the constant threat of irreversible environmental problems.

The government, however, has made it clear that the country will rely on the exploitation of the minerals, and therefore it will continue to be an important issue in Peruvian future.

Aside from mining, is also true that since 1993 the country embarked on a series of economic and political reforms that brought up a sense of stability with no precedence in the republican history. To give an example, the Peruvian central bank has been perhaps the most successful central bank in the region. Since the introduction of the inflation target scheme in 2001, the average inflation rate has been 2.7%.

It is this combination of political and economic upheaval with periods of stability what perhaps makes Peru an interesting case study. Through the chapters in this thesis, I observe the evolution of cities, the behavior of individuals, or the beliefs of adults from 1993 to 2012. More precisely,the first chapter is concerned with the determination of beliefs about trust and identity for a group of individuals I observe from 2007 to 2012. The second chapter proposes a methodology that elicits individual attitudes towards democracy from 2002 to 2012. In the final chapter I study the reaction of local labor markets to a price boom in the mining

industry, the period of analysis is framed by two population censuses: 1993 and 2007.

Yet, the chaotic period which precedes Fujimori's self-inflicted coup d'etat also plays a role among the cities or individuals whose information I observe from 1993 to 2012. Especially in the first two chapters, where I analyze the effect of economic uncertainty experienced by individuals during their past, or their exposure to the conflict that ravaged the country from 1980 to 2000, or the exposure to military autocratic regimes.

1 The Shaping of Trust and Identity After Conflict in Peru

1.1 Introduction

The destruction that ensues after a war is also psychological. Does the psychological trauma have a lasting effect on the individual's beliefs? Can the individual trauma determine the degree of trust that an individual exerts towards public institutions? Or the level of identification with his or her peers?

If violence has an effect on the way people behave, at what time in the course of a lifetime is this determined? Previous works have highlighted the effect of macroeconomic disturbances during a period called the *impressionable years,* on beliefs. The hypothesis about the *impressionable years*, as defined by Giuliano and Spilimbergo (2014) is:

> *"[...] states that core attitudes, beliefs, and values crystallize during a period of great mental plasticity in early adulthood (the so-called impressionable years) and remain largely unaltered thereafter. Evidence of significant socialization has been found between 18 and 25 years of age (Krosnick and Alwin; 1989)"*

Giuliano and Spilimbergo (2014) find that individuals who experienced a recession while young, that is, between 16 and 25 years, believe that success depends more on luck than effort, favor redistribution policies and tend to vote for left-parties.

A natural extension for the Giuliano and Spilimbergo (2014) analysis follows: if recessions, as traumatic as they can be, experienced during the formative years, have a lasting effect on how people think, how a war can influence people's beliefs?

In this chapter I am interested in understanding how conflict affects social capital in the form of trust and identity. I use Peru as my context of analysis. Peruvian data and history offer the possibility to follow different cohorts of individuals who were exposed to various degrees of violence in a geographically heterogeneous setting. Also, the ethnic composition of the country and the unusual cruelty experienced by a particular ethnic group allows to determine whether biased violence destroys any sense of integration or social cohesion. The Peruvian case is also interesting because the violence that was targeted towards the indigenous population was exerted almost in similar proportions by the army and the guerrilla

movement. In that regard, it is interesting to investigate whether the degree of trust towards government institutions was harmed by the government intervention rather than the rebel forces[5].

There is a growing literature that attempts to understand the effect of wars on personal beliefs, which is the main reference for this chapter. In a recent review, Bauer et al. (2016) survey 23 studies on conflict, and argue that individuals exposed to violence tend to increase their social cooperation at the local level, including community participation and pro-social behavior. According to the authors, this may be part of the explanation for quick recovery periods: a sustained sense of cooperation emerging in the aftermath of destruction and destitution. Focusing on trust, Bauer et al. (2016) find that many of the studies surveyed find no effect on trust over other individual's groups, while there is a small positive effect on the trust toward members of the same group. What are the reasons for that behavior? Can we generalize this idea? How does this interact with the internal dynamics of ethnic groups?

In another related work, Adhvaryu and Fenske (2014), find that exposure violent events of African individuals from ages 0 to 14 has little impact on later attitudes and behavior. Particularly on general trust, the authors find no statistically significant effect.

In this chapter, I follow this growing literature on the determination of beliefs and try to understand whether the exposure to violence during the impressionable years period exerted any influence on the way people form their beliefs. With this I aim to contribute to the explanation of the effects of wars on people's beliefs. Additionally, I evaluate whether the interaction of violence exposure with an ethnic dimension helps to shed light on the way war affects beliefs.

From a general perspective, economic literature has long recognized that culture and institutions matters for development, while also been aware that culture ans institutions are generally intertwined. Putnam (1993) for instance sets trust among the pillars from which reciprocity, civil engagement and collective existence unravel to provide modern democracy a framework that leads to economic prosperity. In the same lines, Algan and Cahuc (2010) exploit variation in migrants origins to demonstrate that inherited trust causes growth. From

[5]Using the same empirical strategy and data León (2012) finds that exposure to violence at early childhood reduces the years of schooling by 0.31 for Peruvian individuals.

the view of developing countries trustworthy institutions have been highlighted as n important determinant of prosperity (Bratton and E; 2016).

On the other hand, trust and culture can also be shaped by traumatic experiences. Bauer et al. (2016)find a positive effect of wars on social capital by promoting state formation and nation building through enhanced cooperation after the conflict is over. On the negative side, conflict may switch behavior from honesty to corruption. Empirical evidence shows that conflict increases participation in community organizations and political engagement (Bellows and Miguel; 2006, 2009; Blattman; 2009)

The literature on identity has focused on the counterpoint between national identity and ethnic identity. Masella (2013) for instance, finds that in countries with high ethnic diversity, nationalist feelings are less strong in minority groups than in the majorities, while in countries with low ethnic diversity, the reverse is true. Miguel (2004) estimated that promoting the sense of nation among heterogeneous communities in Tanzania (nation building) allowed this group to achieve better economic outcomes than diverse communities in the Kenyan regions. On the determinants of identity preferences, Eifert et al. (2010) find that closer to an election, individuals in Africa feel more identified with their ethnic roots rather than to occupational group or social class. While the implications of assuming one type of identity over social preferences have been studied by Chen and Li (2009). The authors show that feeling of belonging to a group reduced envy among individuals.

Therefore, this chapter makes three contributions. First, it addresses the impact of armed conflict on beliefs, namely trust and identity, when the violence took place during the impressionable years period of the individuals. Following Giuliano and Spilimbergo (2014), the chapter seeks to understand whether other type of macro shocks at that relevant age (from 16 to 25) have the same relevance in the determination of beliefs. Second, it tries to disentangle whether the identity of the perpetrator sheds some light in the determination of beliefs. And third, it analyses whether the interaction of violence with a measure of ethnicity or race that reflects an important dimension of the population of the study context, offers an alternative explanation.

Methodologically, in this chapter I exploit the variation in conflict location and birth cohorts to identify the effects of conflict on two types of social capital traits: trust and identity.

I differentiate the effect by the identity of the perpetrator, and by the indigenous origin of the individual.

To preview the results, I show that Peruvian individuals exposed to violent events during their impressionable years trust less government institutions, and feel less identified with their neighbors, while more identified with religious groups. The estimated effect is small and heterogeneous depending on the identity of the perpetrator. The effect on identity is also heterogeneous depending on the degree of connection to the indigenous dimension of the individual. In concrete, individuals who own a an agricultural plot which is embedded in a local institutional setting that proxies indigenous arrangements, exhibit a smaller degree of identification with locals (or neighbors), while higher levels of identification with ethnic and race groups. This last results is better understood in the historical context of Peru, where the conflict experienced between 1980 and 2000 had a decisively political motivation which at the same time disregarded the ethnic component presents in the country since the times of independence. The systematic attack from the army and the guerrilla eroded the identification of the individuals with their locals, while deepened the identification with their ethnic peers. Together these results suggest that wars that reached individuals during their impressionable years can have lasting effects on the trust they exhibit, but overall, on the sense of identity.

The chapter is organized as follows: section 1.2 discusses the data used in the analysis, and overviews the historical relevance of the conflict period and the ethnic and race dimension through the discussion the the communal land setting and its interaction with native and peasants. Section 1.4 presents the empirical framework used to identify the effect of conflict. Section 1.5 discusses the main results for the effect of conflict during the impressionable years. Section 1.6 outlines possible mechanisms and discusses some avenues for further research. Finally, section 3.6 concludes.

1.2 Historical Overview

In this section I provide the historical background for the armed conflict and the ethnic dimension.

1.2.1 The Armed Conflict

"People were afraid, and shouted: 'They are coming! They are coming! If not the Shining Path, the soldiers![6]'

From 1980 to 2000, Peru experienced the longest and bloodiest conflict in its republican history. The number of Peruvians who perished victims of it was approximately 69,280[7] according to the estimates of the Commission for the Truth and Reconciliation (CVR for its acronym in Spanish, CVR 2003[8]). A number that surpasses in magnitude the number of Peruvians dead in the other two major wars, the war against Chile in 1879 and the independence war in 1821.

Unlike many of the conflicts surveyed in Bauer et al. (2016), the driving force in the Peruvian conflict did not have an ethnic origin, but a political motivation. The two forces confronted were the Peruvian Communist Party (PCP), led by its more radical members grouped in what ended up being *Sendero Luminoso* (the Shining Path, SL in Spanish), and the government (who was also supported by local self-defense committees). A second faction of rebels joined the conflict in 1984 under the name of Tupac Amaru Revolutionary Movement (MRTA in Spanish).

The estimates of the CVR attribute 54% of the casualties to the PCP-SL, 43.5% to the government and 1.5% to MRTA. But perhaps, the most striking fact that complements the almost even distribution of victims between PCP-SL and the government, is the proportion of victims whose origin was indigenous: 75% of all victims had *Quechua*[9] as their mother tongue. This percentage is an indication of the degree of violence biased towards indigenous population, considering the percentage of people whose main language is of any ethnic origins.

Partly, the large percentage of native victims is explained by the high proportion of native speakers in the highlands. However, the 75% of *Quechua* victims, is even higher than the proportion of Quechua speakers among the Ayacucho (the region with the highest toll) population in 1994, 70.6%. Numerous testimonies collected by the CVR provide a detailed

[6] Testimony N 450066. CVR (2003)

[7] Section 1.3 provides details about the estimation procedure used by the CVR.

[8] Available online at: http://cverdad.org.pe/ifinal/

[9] The 1994 population census reported that 19.42% of the population had any native language as mother tongue, out of which *Quechua* was 16.46%, and *Aymara* was was 2.28%. The population 2007 census reported 15.68% of the population as native speakers, and 13.02% in particular as *Quechua* speakers.

account of the extraordinary degree of cruelty exerted from both, the PCP-SL and the army, towards this group of Peruvians. It was common that some communities experienced attacks and counter-attacks from both group after the other. Examples of this are the massacres in the towns of Lucanamarca and Accomarca.

On the 22th of March 1983, the self-defense committee from the community of Lucanamarca, in Ayacucho, captured and executed a commander from PCP-SL. On the 2nd April, 60 members of PCP-SL entered the community of Lucanamarca and executed 69 inhabitants. The order from the leaders of PCP-SL was to be as cruel as possible to provide an exemplary demonstration of the consequences of opposing the revolution. Indeed, the accounts of the events elaborated by the CVR indcate that the victims were killed with axes and machetes. 18 victims were children whose slaughtered bodies were exposed in the main plaza. Some victims died after being submerged in boiling water. The leader of PCP-SL, Abimael Guzman, justified the action as a vindication of the revolution.

The other event, in Accomarca (also in Ayacucho), was conducted by the Army. On the 14th of August 1985, a battalion from the Peruvian army entered the community of Acomarca, which was under the suspicion of being a base for PCP-SL. After gathering the local population in the main plaza of the locality, 30 children, 27 women and 12 men were brutally murdered. After shooting everyone and after having raped the women, the army locked the three groups of victims inside some houses and burned the entire community.

These two events illustrate the cruelty of the conflict, which did not follow any war convention, over the indigenous population. The CVR signals this an evidence of the ethnic fragmentation that added a never seen degree of violence in the country. The same ethnic fragmentation, however, has been pointed by the CVR as the reason to explain the failure of PCP-SL's revolution. PCP-SL's political conception and strategy were at odds with the cultural and social characteristics of the peasants communities and populations affected by the violence. The rebel group adopted and attempted to implement a philosophy based on China's cultural revolution which relied on a vertical imposition of the mandates of the party over the dominant class in favor of the peasant class. This classification was not so evident in the communities of the highlands, where the social interaction within the communities followed an historical pattern of cooperation within the community with little difference

among its members. If anything, the terrorist actions provided the opportunity for old confrontations between communities to become violent.

When the government allowed the army to take control of the situation, the ethnic differences became also relevant. The CVR has documented several case studies which reveal an army exerting a similar degree of disregard for the indigenous population, as the PCP-SL. However, in the evolution of the conflict, the army was quicker than PCP-SL to gain cooperation from the local communities, which ultimately turned out to be vital for the final defeat of the terrorist group.

Chronology

The conflict started officially with PCP-SL declaring war on the Peruvian government. One symbolic action characterizes this declaration of war: on the 17th of May of 1980, members of the terrorist group burnt electoral ballots in the Chuschi district, in the region of Ayacucho. 1980 signaled the restitution of democracy in Peru after 12 years of autocracy under the rule of generals Juan Velasco Alvarado and Francisco Moralez Bermudez. It was precisely general Moralez Bermudez who decided to call for elections in 1980.

It has been pointed out by the CVR that this immediate context explains the rapid expansion of the terrorist actions in the first part of the decade of 1980. Fernando Belaunde Terry won the elections that year, which symbolized the return to democracy after the coup d'etat in 1968. The re-appointed president, therefore, had little incentives to coordinate with the army for a stronger response to these initial actions of the PCP-SL. On their part, the PCP-SL's decision to embark into the armed conflict came as a consequence of the radicalization of the protests led by the left parties during great part of the 12 years of military dictatorship. This, coupled with what was the beginning of a long period of economic turmoil, facilitated the expansion of the terrorist's actions.

The decision of PCP-SL to start their operations in Ayacucho is related to a growing discontent among educated classes in that region. As explained by CVR (2003) and León (2012), the rapid expansion of access to education among population during the mid 70's did not find a match in the labor market. Therefore, a growing number of dissatisfied university and secondary education students found the ideas of PCP-SL appealing under that context. Some

universities and schools in Ayacucho had a reputation for leading movements in favor of the expansion of education. The PCP-SL leaders took University San Antonio de Huamanga as their basis for the indoctrination phase.

Figure 1.1 plots the annual evolution of the number of events and victims for the period between 1980 and 2000. The number of events peaked twice, in 1984 and 1989, while the number of victims in one year reached its maximum level in 1984. Geographically, the conflict started in the region of Ayacucho, but at some point it reached almost all the country. 1.2 plots the geographical dispersion in 1980, at the beginning, and in 1989 at its peak. By 1989, there were violent actions even in the far north of the country.

According to the CVR, the conflict can be classified in 5 periods:

1. ***Beginning of the armed conflict (May 1980-December 1982)***. PCP-SL burns electoral ballots in the district of Chuschi, in the Ayacucho region. Little press coverage of the events, mainly explained by the interest on the transition to the new democratic government. Additionally, the police institution was undergoing a major reformation to join in one body the three traditional types of police institutions, poorly articulated in the country: Civil Police, Republican Police and Investigation Police. On the 30th of December of 1982, the president allowed the army to take control of the situation.

2. ***The militarization of the conflict. (January 1983-June 1986)***. The army starts to take control of some regions where PCP-SL operated. In those regions the government declared the *State of Emergency*. Both, the army and the terrorists are involved in local massacres, and this marks the beginning a harsh period in the highlands. MRTA joins the conflict in the region of San Martin, with connections to the drug trafficking. 1984 marks the first peak in the number of victims (see figure 1.1).

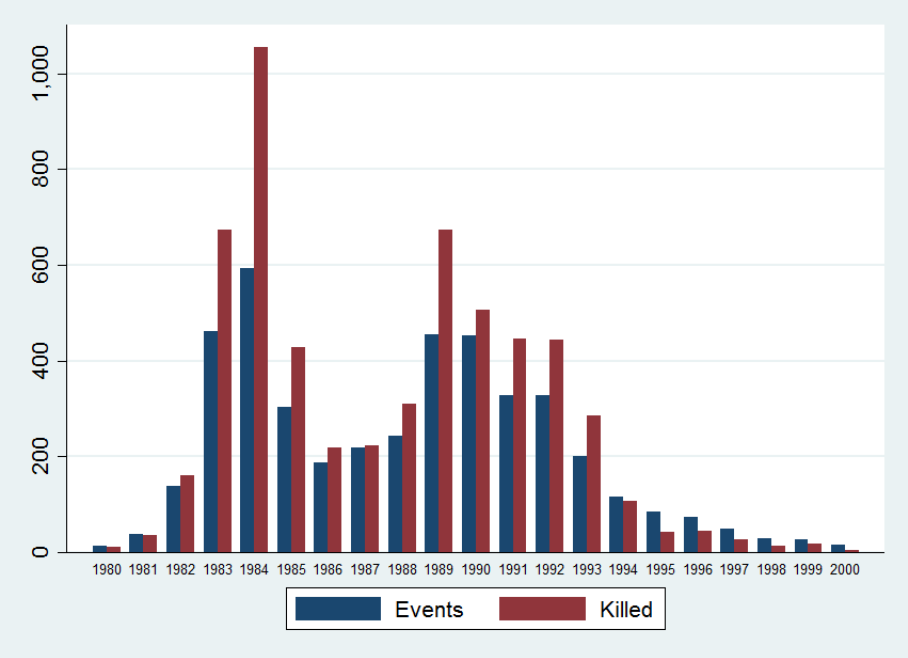
3. ***Violence's upsurge (June 1986-March 1989)***. this period witnesses the opening of new combat fronts in urban areas. Up to 1986 the majority of the violent action took place in rural areas in the highlands. Both terrorist groups liaise with drug cartels, specially in the area where MRTA was located. The national economy was in crisis: Peru started to experience a period of hyperinflation. And the government lost control over the army.

4. ***The peak of the crisis (March 1989-September 1992)***. 1989 registers the second peak in

the number of victims. The army led several counterattacks in the highland region. Huge political instability in the central government: on the 5th of April, 1992, president Fujimori dissolved the parliament, which marked the beginning of his autocratic regime. The leader of PCP-SL, Abimael Guzman, was captured.

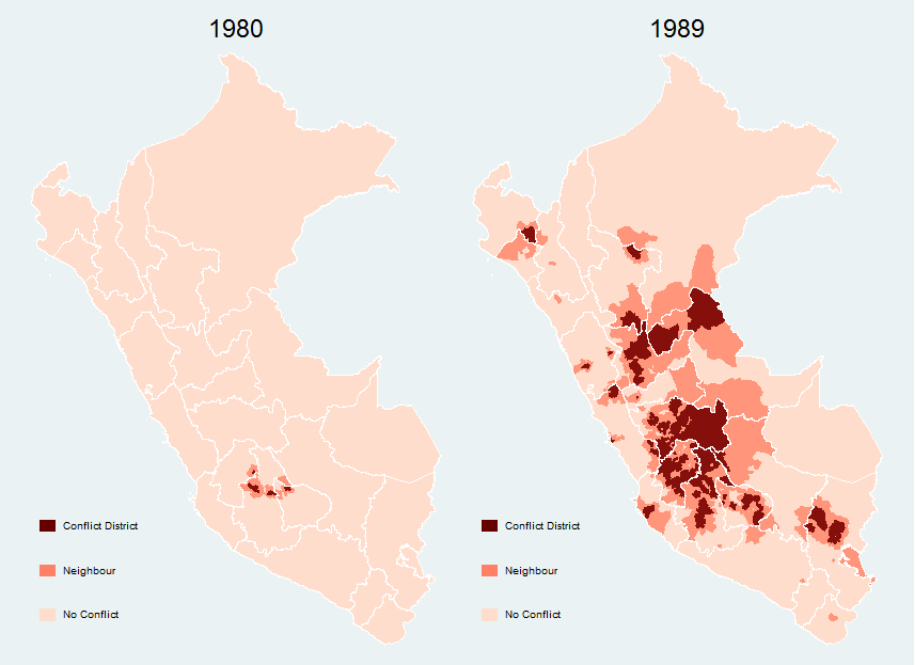
5. ***Decline of violence (September 1992-November 2000)***. After the capture of Abimael Guzman, PCP-SL lost control of many regions. The army, supported by local self-defense committees regained the control of those regions. The government enacted a series of laws to provide amnesty to repented terrorists in exchange for intelligence collaboration.

Figure 1.1: Number of Events and Killed People



Notes: [1] Data source: CVR (2003)

Figure 1.2: Geographical Dispersion of Conflict: 1980 and 1989



Notes: [1] Data source: CVR (2003)

1.2.2 Communal Land, Native Origin and *Comuneros*

The large number of indigenous victims highlights the ethnic bias for the violence exerted during the conflict. It is remarkable that both, PCP-SL and the army punished the indigenous groups with similar levels of cruelty. In that regard, because indigenous groups suffered a disproportionate effect of the conflict, the analysis design should help to understand whether this group reacted differently after the violence. If there was any effect of violence on trust and identification, would indigenous individuals express an heterogeneous response? If there is a type of social organization which groups the majority of native and indigenous population in Peru that is the communal land organization. A complete understanding of the interaction between the ethnic dimension and conflict has necessarily to provide an understanding of the communal land associations and how they relate to the native and individual dimension.

To the best of my knowledge, there are no studies that address this interaction. From the economic point of view, however, some studies have pointed out that there is at least an insurance justification for the probability of engaging in violent activities among members

of the communal land organizations. Guardado (2015) suggests that in the face of coffee price shocks, individuals are less likely to engage in violent activities if they live in a city where there is a high proportion of agricultural land under communal tenancy. The reason for that is that the provision of communal insurance prevents the upsurge of violence. In the explanation of this result, there is an natural connection to the indigenous or ethnic origin of the individuals.

Historical communal setting in Peru have been related to some sort of cooperative behavior, which ultimately signals a certain degree of trust within the community, and in terms of the analysis of Bauer et al. (2016) a more pro-social behavior. It also signals to a degree of identification within the community: if members of the communal association are tightly connected, it is probably that the internal network for risk-sharing is stronger. Following Guardado (2015), the basic measure for this dimension is the importance of the communal setting in the city where the individual grew up. However, this can also be measured at an individual level through the native origin of the adult or his or her individual connection to the communal setting.

In the empirical section I describe three ways, mutually linked, in which I measure the ethnic dimension. The first is an aggregate measure at the district level (as Guardado; 2015): the proportion of agricultural land under the tenancy of peasant or indigenous communities, which I term as communal land. The second and third are individual indicators I obtain from National Household Survey (ENAHO in Spanish): the native origin of the individual, or whether the individual has a an agricultural plot which belongs to a communal association. I refer to the native origin simply as "native" to denote an individual whose main language is either *Quechua* or *Aymara*. While for the last measure, I define an individual as *comunero:* in Peru, this is the terminology to identify a peasant who is part of an indigenous community or conducts agricultural operations within a communal land setting.

In this section I discuss the historical background for the ethnic dimension that permeates these three measures through the explanation of the origin of the communal land associations, and leave a detailed statistical description to section 1.3.

Brief Historical Background on communal land and native origins

Being of native origins in Peru implies some degree of connection not only to the native languages, but also to certain ancient practices. In the highlands of Peru the property of land in hands of indigenous communities has been a matter of long standing analysis. As pointed out systematically, the organization of the communal geographical space has been historically associated with low levels of agricultural production (CEPES; 2005), and more importantly, with the prevalence of indigenous population (Del Castillo; 1997). Therefore, the geographical unit, an agricultural plot for instance, known as communal land, has been inherently linked to individuals speaking a native language, and more precisely, an individual who is part of an indigenous community or conducts his agricultural activities within a communal setting is know as *comunero*. Therefore, the discussion of the native or *comunero* status is implicitly embedded in the discussion about communal land.

Historical accounts more or less date the origin of nowadays associative rural communities, the closest definition of communal land associations, to the colonial period of Peru (Del Castillo; 1992, Del Castillo; 1997, CEPES; 2005, Caballero and Alvarez (1981), INEI (1998)). They were the output of the *Reducciones de Indios*, a term that can be translated as Indigenous Communities. The main purpose of these *Reducciones* was to facilitate the evangelization of the remaining indigenous population after the consolidation of the Spanish conquest. Viceroy Toledo is believed to have started the process by 1570. In doing so, he relied on an ancient, pre-inca type of population organization: the *Ayllu*.

Although the Inca organization was also based on the *Ayllu*, this type of societal arrangement was not an Inca innovation. Incas were known for assimilating good practices from the diverse cultures they conquered, and *Ayllus* were one of them. Therefore, *Ayllus* pre-date the consolidation of the Inca empire.

An *Ayllu* was basically a family organization, comprised of one or many families, in which a *Curaca* was the visible head of the extended family. In the organization of the *Ayllu* all members had to provide labor in equal parts and share the produce of the land. As it became patent during the Inca period, the main purpose of the *Ayllu* was to provide food to all members. But the activities of the *Ayllu* were not constrained to the core group of families, it also had to provide food and labor to the state (the Inca). In retribution, the *Ayllu* would

receive food from the state in case of shortage. Each Ayllu offered three forms of communal contribution, the *Mita*, *Minka* and *Ayni*.

The *Mita* was a system through which members of the *Ayllu* delivered labor when the central government (the Inca) required the construction of large scale infrastructure like temples, bridges, or mining exploitation, among other labor intensive activities. The arrival of the Spanish colonizers precisely took advantage of this organization to exploit a mercury mine in the central highlands, that depleted most of the male population from the *Ayllus* close to the mine, which also had persistent negative effects until today, as documented by Dell (2010).

The other two forms of labor, *Minka* and *Ayni*, represented also a type of communal labor arrangement, but in favor of the *Ayllu* itself. In particular, the *Minka* system sought to provide a source of communal labor for the construction of small infrastructure in favor of the local *Ayllu*. While the *Ayni* was a type of risk-sharing mechanism through which a family within the *Ayllu* would support another family from the *Ayllu* during the bad times.

The colonial period of Peru that followed the arrival of the Spanish conquerors kept this types of organizations and used some of their labor arrangements. After the independence war, and by the beginning of the republic period in 1821, *libertador* Simón Bolívar sought to disintegrate the inherited *Reducciones* because he thought they were a colonial institution that prevented the integration of the indigenous population into the new society.

History denied this wish to the *libertador*, and associative rural communities remained informally alive until 1920; year in which the enactment of a new constitution recognized the existence of these communities. That was the first year the Peruvian government promulgated a law aimed to recognize and protect them.

In the following years, the promulgation of three new constitutions, an agrarian reform, and the intense guerrilla period under analysis, did not change the core of the legislation around the communities, neither the *de-facto* structure: the rural communities survived until today. Several modifications to the law were introduced during this period. Among the most important, in 1992: members of the community can sell the land if they want. Yet, the agricultural census of 2012 reports that 60.7% of the agricultural land in Peru is still held by these communities.

Interestingly, these communities are still engaged in the ancient practices documented during the colonial and Inca period. Data from the 2012 agricultural census indicate that 59.13% of the rural communities still engage in the *Ayni*, 56.57% in the *Minka*, 0.92% in the *Mita*, and 83.78% in any of the three. Also interesting is the little cooperation between communities. Despite the fact that communities are still engaged in cooperation activities within the community, the interaction between them is very limited. The agricultural census of 2012 reports that the percentage of communities engaged in multi-communal businesses is 3%. This fact is not trivial given the context of the guerrilla period. Previous sections noted that the violence which ensued after PCP-SL's arrival to the highlands had some basis on the historical conflicts between communities.

The CVR conceives the interaction of the ethnic component with the presence of the guerrilla or the army as a key factor explaining the unparalleled cruelty in Peruvian history that this conflict represented.

1.3 Data

I use three data sources for the empirical analysis. With the National Household Survey I measure individual's beliefs, demographics as well as their individual connection to communal plots. With the agricultural census of 1994 I measure the share of agricultural land under communal tenancy. Finally, the CVR is the source for the conflict data at district level.

Although the formation of beliefs during the impressionable years period is the main analysis of this chapter, I start with the description of the data used to measure the exposure to conflict during the impressionable years, as well as the measures for the ethnic dimension.

1.3.1 Conflict

Based on the reports for approximately 24,000 victims recorded in different sources, the CVR was able to extrapolate the number of victims to 69,280. The 95% confidence interval for this estimation was [61,007-77,522]. The method utilized for the estimation of the number of victims is called *Multiple System Estimation* (MSE).

The MSE methodology has been recently used in the estimation casualties for different

conflicts[10]. The original version was developed by a Danish scientist, Johannes Petersen in 1860 to calculate the number of fishes in a pond[11]. The most recent applications, however, use the methodology to estimate the population size of regions where census data collection may be inaccurate due to isolation. In the conflict context, Manrique-Vallier et al. (2013) provides a review of the application of the methodology to conflicts in Guatemala, Kosovo, and Peru.

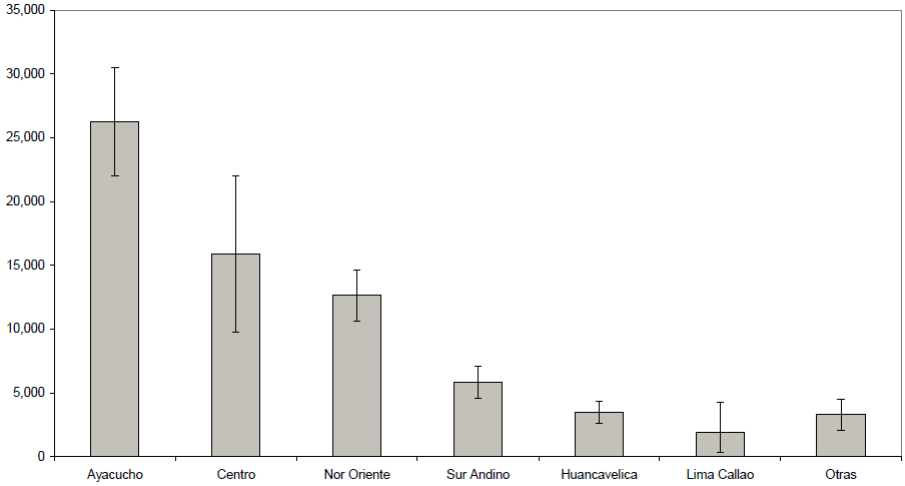
The original implementation of the methodology serves as an illustration of its procedure. If we wanted to know the number of fishes in a pond, we could proceed to catch a certain number, say 100. After marking the 100 fishes we return them to the pond. On a second catch of again 100 fishes, we count how many of the initial 100 were caught again. Imagine this number as 20. Therefore, the whole population fishes in the pond should be 500. The idea for the estimation is simple. If the two random catches are independent and every fish has the same probability of being caught, the ratio of the number of fishes caught in the second catch over the size of the first catch (20/100), should be proportional to the size of the second catch over the total number of fishes that populate the pond (100/N). After a simple manipulation of these two proportions, the total number of fishes is estimated as 500. Appendix A provides more detail on the methodology used by the CVR to produce the estimates.

Figure 1.2 shows that at its second peak in 1989, the conflict was widely widespread, reaching areas beyond the geographical origin, Ayacucho (left panel). Figure 1.3 plots the regional estimates. Ayacucho, concentrated 26,259 casualties, which represents 38% of the total number of victims. While 21.7% of the casualties were registered in the central highlands region. The Nor Oriente region refers the jungle and around 18% of the casualties were reported from that region. Sur Andino refers to the southern highlands, and approximately 8% of the victims were from this region. The capital Lima, as well as Callao, its harbor, and the rest of the country experienced smaller intensity in the number of victims.

[10]However, it has been also used in other contexts of public policy sensitivity like drug use. See King et al. (2013)

[11]Although the origin of the methodology is disputed. See Goudie and Goudie (2007)

Figure 1.3: Conflict: Geographical Intensity



Notes: [1] Figure taken from the CVR Final Report. CVR (2003), annex 2. [2] The grey bars indicate the average estimate while the spikes are the 95% confidence interval estimatios.

Following Giuliano and Spilimbergo (2014), I define the individual's impressionable years the 10 years period when the individual's age ranges between 16 and 25. The city of reference for this period is the district of birth. Hence, I measure the number of impressionable years that the individual was exposed to violence. Measuring the exposure to violence this way, follows the approach used by León (2012), where the author measured the exposure to violence and its impact on schooling attainment at the individual level. Therefore, I define the variable $killed_{i,ro}$ as:

$$killed_{i,ro} = \sum_{t=16}^{25} violence_{ro,t} \tag{1.1}$$

Where $violence_{ro,t}$ is a dummy variable that indicates that in year t, in district ro, there were at least one killed person due to the violence and it coincided with any impressionable year of the individual. Therefore, $killed_{i,ro}$ measures the number of impressionable years that the individual was exposed to conflict in their district of birth. Table 1.1 summarizes the main explanatory variable. For the national sample of individuals in ENAHO from 2007 to 2012, 90.54% did not experience any type of violence at their district of birth during their impressionable years. 9.46% of the individuals had at least one impressionable year affected by conflict at their district of birth. On average, an individual experienced 0.23 years of violence that overlapped with their impressionable period. The standard deviation for this

mean is 0.95.

A variation of this definition considers army and terrorist violence separately. The definition would be the same, but the violence count considers only violence attributed to the army and police forces on one side, and the terrorists on the other. The last two columns of table 1.1 show the frequency count for those variables.[12] On average individuals experienced more terrorist violence than army violence during their impressionable years. On average individuals experienced 0.16 years of terrorist violence during their impressionable years period, which is almost the double compared to the average number of impressionable years affected by army violence, 0.07.

Table 1.1: Number of Individuals With Impressionable Years Affected by Conflict

	Total		By Group	
N. of Years	Freq.	(%)	Army	Terrorist
0	142,301	90.54	150,988	146,206
1	7,058	4.49	4,076	5,540
2	3,439	2.19	577	2,138
3	1,195	0.76	541	1,130
4	1,008	0.64	597	634
5	694	0.44	156	444
6	393	0.25	88	384
7	338	0.22	55	376
8	359	0.23	64	106
9	239	0.15	25	84
10	143	0.09	0	125
Mean	0.23		0.07	0.16
S.D	0.95		0.47	0.79

Notes: [1] Data sources: ENAHO 2007-2012 and CVR (2003). [2] Impressionable years is the period of an individual between the age of 16 and 25.

1.3.2 Communal land, native or *comunero*

The three measures that I use to approach the ethnic dimension in the analysis come from different sources. The first, the share of communal land in the districts, is an aggregate measure I calculate using information from the 1994 agricultural census. The other two measures, the status of native speaker, or the status of *comunero*, are measures at the individual level that I retrieve using information from ENAHO itself. In this section I present some descriptive statistics to to explicitly reveal the positive correlation between these measures.

[12]Government violence accounts for killings associated either to the army or the police, according to the accounts of the CVR (2003), while terrorist violence groups killings associated to either the Shining Path or the MRTA

In the sample from ENAHO I use, 28% of the individuals are characterized as native, while 5% are *comuneros*. The data source to estimate the district share of communal land is the Agricultural Census of 1994. For a district to be considered a communal district, the share of agricultural land under the tenancy of peasant or indigenous individuals has to be larger than 0.5.

Table 1.2: Land Arrangement (Thousands of Ha.)

	1994	2012	Change
Total Land	**35245.8**	**36426.5**	**3.35%**
Individual Land	14027.2	13250.1	-5.54%
(%)	39.80	36.37	
Land in society	1242.5	613.4	-50.63%
(%)	3.53	1.68	
Coop Land	355.0	44.9	-87.37%
(%)	1.01	0.12	
Communal Land	14089.1	15515.3	10.12%
(%)	39.97	42.59	
Native Land	5251.9	6587.7	25.43%
(%)	14.90	18.08	
Other type of Land	280.1	415.2	48.23%
(%)	0.79	1.14	
Average Share by District	0.42	0.37	
Average Share by Province	0.48	0.49	
Number of Districts	1801	1837	
Number of Provinces	195	195	

Notes: [1] Data source: CENAGRO 1994 and 2012. [2] Communal land in 1994 includes the after agrarian reform farmers' groups. Native land is the denomination of communal land from the forest. [3] Share is the sum of communal and native land over the total land surface in the district. When missing (mainly metropolitan Lima), it was replaced by zero.

I start by describing the persistence of the communal land setting between the two agricultural censuses: 1994 and 2012. I focus on tenancy at the district level, which is the variable I will use to differentiate a district as a communal district. Table 1.2 reports the composition of agricultural land by the type of tenancy recorded in the agricultural censuses of 1994 and 2012. Total agricultural land in Peru grew merely by 3.35% between 18 years [13]. Communal land, the main concern of this section, is the land in hands of indigenous population. A community is able to register this type of settlement, which constitutes a type of private ownership, although manged in a communal setting. The government decided to keep the *communal* denomination to refer to indigenous population from the highlands with *Quechua* or *Aymara* origins, and coined a new term, *native*, to refer to the indigenous population in

[13]The previous agricultural census of 1974 reported a total agricultural surface of 23.545 millions of Ha. In consequence by 1994 the total agricultural land grew by 49.7% by a similar period.

the east side of the Andes, the forest. Combined, the land in hands of these two groups was 54.9% of the total land in 1994, and 60.7% in 2012. Hereafter, unless specified, I will refer as *communal* to the combination of these two types of land ownership.

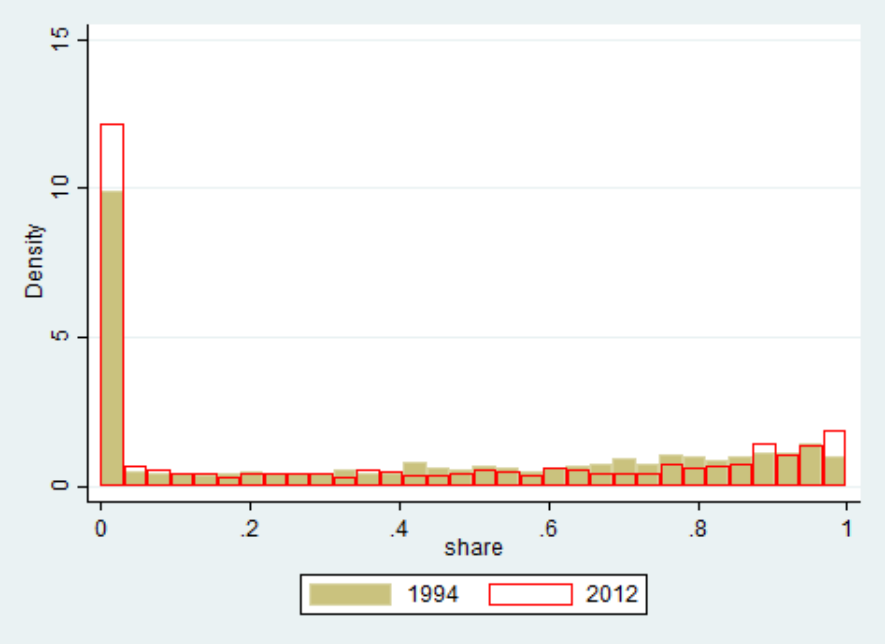
The land owned by individuals remained as the second largest type of tenancy. It represented 39.8% of total land in 1994 and 36.37% in 2012. This smaller share is explained by a reduction, 5.54%, in the land in this category. Society and cooperative ownership were the categories with the largest drop in their share: 50.63% the first, 87.37% the second. This is a consequence of the adjustments after the deactivation of the agrarian reform in 1982[14].

The average share of communal land in a district was 0.42 in 1994, and 0.37 in 2012. However if the unit of observation is the province, such shares are 0.48 in 1994 and 0.49 in 2012. Figure 1.4 plots the distribution of the district share of communal land for both periods: it is clear that such distribution has changed little from 1994 to 2012. Figure 1.5 plots the correlation between the share of communal land in 1994 and the share of communal land in 2012 at the district level. The correlation between these two periods is 0.63[15].

[14]General Juan Velazco launched an agrarian reform in 1969 that expropriated large extensions of land from landlords to be transferred to local farmers. This reallocation of land did not touch large communal territories and mainly affected private territories in the northern coast of Peru. By 1982, a new military government annulled the reform but did not seek the restitution of previous ownership.

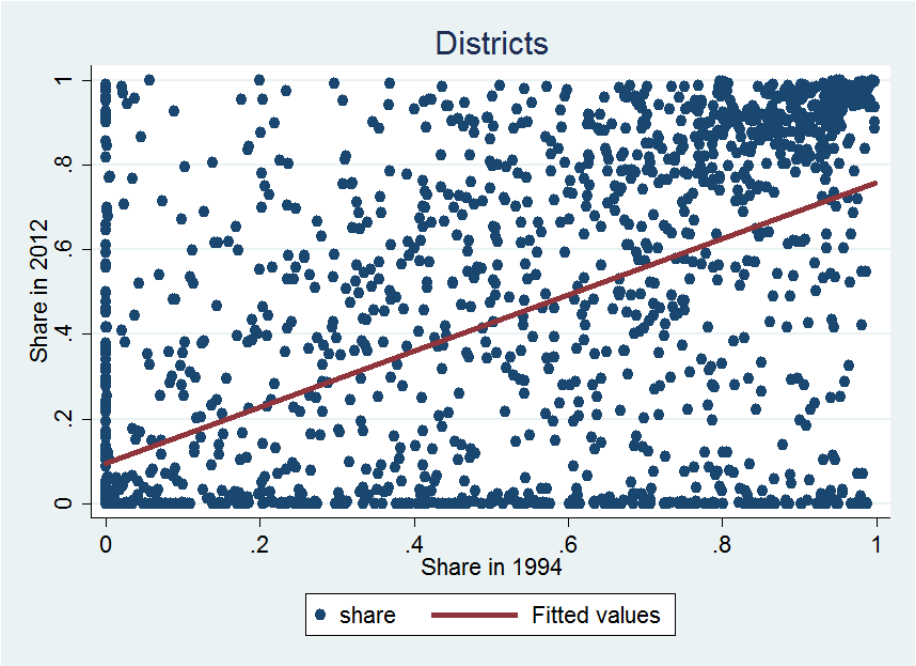
[15]At at bigger administrative level, provincial, the correlation coefficient between 2012 and 1994 share of communal land is 0.87.

Figure 1.4: Share of District Communal Land



Notes: [1] Data source: CENAGRO 1994-2012. [2] Histograms plot the district share using the 1993 border definition, which consisted of 1791 districts.

Figure 1.5: Correlation Share of District Communal Land: 2012 and 1994



Notes: [1] Data source: CENAGRO 1994-2012. [2] The estimation uses 1791 districts each period, according to the 1993 border definition.

Not I turn to the demographic differences. Table 1.3 provides statistics at the individual level broken down by the two types of districts the individuals live in: communal and

Table 1.3: Summary Stats: Demographics at Individual Level

	Non-Communal	Communal	Total
Years of schooling	9.401	7.985	8.942
	(4.429)	(4.706)	(4.569)
Male	0.504	0.529	0.512
	(0.500)	(0.499)	(0.500)
Native	0.114	0.446	0.222
	(0.318)	(0.497)	(0.415)
Comunero	0.0227	0.0789	0.0410
	(0.149)	(0.270)	(0.198)
Coast	0.476	0.266	0.408
	(0.499)	(0.442)	(0.491)
Age	41.20	44.07	42.13
	(15.97)	(16.51)	(16.20)
Observations	104410		

Notes: [1] Data source: ENAHO 2007-2012. [2] Unweighted averages. Standard deviations shown in parenthesis. [3] The sample includes individuals at least 18 years old.

non-communal. It immediately reveals a positive association between the three variables: both, the proportion of native speakers or *comuneros* are high in communal districts. The proportion of native speakers in communal districts is 0.446 while in non-communal districts is 0.114. The reason for this is that the definition of communal for a district uses the 0.5 threshold in the share of agricultural land under the tenancy of communal associations. It is of course possible that there are native individuals in districts with communal land share smaller than 0.5. Similarly, the proportion of *comuneros* is 7.9% in communal districts while it is 2.3% in non-communal districts. On this number, first it is worth pointing out the relatively low rate of *comuneros* nationwide, 4.1% of the sampled individuals. A possible explanation for this the fact that although some individuals live and work in a communal land setting, they conduct activities in separate plots that the community granted from the community authorities, which ultimately represents a type of private ownership in their view. This situation is discussed in Del Castillo (1992) and Del Castillo (1997). This table offers additional information on the communal districts at the individual level. They have less educated people: on average an individual living in a communal districts has 7.99 years of schooling, while individuals in non-communal districts accumulate 9.4 years of schooling on average. Communal districts are slightly more populated by men, and they are not located in the coast. The average age of an individual living in a communal district is 44 years while

individuals from non-communal districts are on average 41.2 years old.

Table 1.4: Demographic Variables at District Level, broken down by district's share of communal land (1994 and 2007/2012)

	1994			2007/12		
	$<= 0.5$	> 0.5	Diff.	$<= 0.5$	> 0.5	Diff.
Quechua	0.208	0.426	-0.218***	0.229	0.403	-0.174***
Aymara	0.028	0.022	0.005	0.017	0.028	-0.010
Other Native	0.004	0.015	-0.011***	0.003	0.023	-0.020***
Spanish	0.760	0.537	0.223***	0.748	0.543	0.205***
Population (Ln.)	7.827	7.036	0.791***	8.231	7.703	0.528***
Age	34.893	36.128	-1.235***	39.873	41.147	-1.275***
Male	0.559	0.577	-0.018***	0.508	0.505	0.002
High Skilled	0.115	0.082	0.033***	0.176	0.133	0.043***
Agriculture Labor	0.549	0.635	-0.086***	0.547	0.618	-0.070***
Events per 10,000 pers.	5.431	16.044	-10.614***	2.992	5.510	-2.518**
Victims per 10,000 pers.	93.014	679.228	-586.214*	39.584	353.834	-314.250*
Victims (Army) per 10,000 pers.	29.472	435.628	-406.156	11.502	261.938	-250.436
Victims (Terrorists) per 10,000 pers.	61.839	220.393	-158.554*	24.320	87.671	-63.351*
Observations	1791			1791		

Notes: [1] Data sources: population census 1994 and 2007 for demographic variables. Agricultural census 1994 and 2012 for share of communal land. [2] Districts in 2012 use 1994 borders. [3] All demographic estimations refer to population aged 18 years or more. [4] Differences that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

In table 1.4 I compare district averages using the agricultural census and the population census. I match information of the 1994 agricultural census with information of the 1994 population census; and information of the 2012 agricultural census with data from the 2007 population census. The idea is to determine whether 1994 differences in population variables between communal and non-communal districts are still relevant in the period 2007-2012. The general conclusion is that although 14 years separate the population censuses, the demographic differences have remained.

The detail of such differences as follows: table 1.4 shows that in both rounds of agricultural censuses districts with a high share of communal land show a larger prevalence of population speaking *Quechua*, but not *Aymara* neither in 1994 or 2007. The proportion of people speaking other native language is again higher in communal districts for both periods. Not surprisingly, the proportion of Spanish speakers is higher in the non-communal districts for both periods. Communal districts are populated by relatively elderly adults. Communal districts are less populated than non-communal ones. The proportion of males in the communal districts was slightly (and statistically significant) higher in communal districts. However there are no differences in the 2007 data. The proportion of high-skilled workers

(with at least technical education) was smaller in the communal districts in both periods. The proportion of agricultural labor force was also higher in communal districts in both periods.

The last four rows of table 1.4 provide an insight of the intensity of violence experienced by this districts. The way these four variables are measured consider the number of events, victims (overall, attributed to the army and to the terrorists) for the whole period of 1980-2000 scaled by 1994 district population. The comparison in the first two columns classifies districts as communal or not depending on the communal land share estimated with the 1994 agricultural census; while the second comparison uses the 2012 agricultural census to classify the districts. The only difference that is statistically significant is the number of events per 10,000 people. It was higher in communal districts. The number of victims scaled by population in the communal districts is larger in magnitude than the victims per population in the non-communal districts, however it is only for the difference in terrorist's victims that the difference is statistically significant at 10%.

Table 1.5 replicates the exercise but using the proportion of native speakers in the district. Something that emerges clearer in this table is the higher number of casualties and events in districts with a higher proportion of native speakers.

Table 1.5: Demographic Variables at District Level, broken down by district's share of native language (1994 and 2007)

	1994			2007		
	$<= 0.5$	> 0.5	Diff.	$<= 0.5$	> 0.5	Diff.
Quechua	0.074	0.747	-0.673***	0.073	0.759	-0.686***
Aymara	0.010	0.055	-0.045***	0.008	0.050	-0.042***
Other Native	0.006	0.015	-0.010**	0.007	0.021	-0.014***
Population (Ln.)	7.685	7.061	0.624***	8.140	7.775	0.365***
Age	35.714	34.952	0.762***	40.069	41.022	-0.953***
Male	0.555	0.591	-0.035***	0.514	0.492	0.022***
High Skilled	0.115	0.071	0.044***	0.187	0.102	0.084***
Agriculture Labor	0.553	0.654	-0.101***	0.539	0.649	-0.109***
Events per 10,000 pers.	2.232	25.396	-23.165***	1.361	9.382	-8.021***
Victims per 10,000 pers.	17.279	1004.155	-986.876***	10.810	480.313	-469.503***
Victims (Army) per 10,000 pers.	2.701	611.969	-609.268**	1.647	336.048	-334.401*
Victims (Terrorists) per 10,000 pers.	13.434	361.271	-347.837***	8.652	133.304	-124.652***
Observations	1791			1791		

Notes: [1] Data sources: population census 1994 and 2007 for demographic variables. [2] Districts in 2007 use 1994 borders. [3] All demographic estimations refer to population aged 18 years or more. [4] Differences that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

In sum, what all these statistics show is that communal districts were inherently different than non-communal ones, both, in terms of their demography and the violence they were

exposed. And such differences have remained through time. The empirical section will take advantage of this characteristic and define the 0.5 threshold using the 1994 distribution. Classifying districts depending on the proportion fo native speakers provides a similar picture where the intensity of violence is higher for native districts.

1.3.3 Individual's beliefs (and demographics)

For the variables about trust and identity, I use the National Household Survey (ENAHO) for the period 2007-2012. The survey collects information about the household and the individual that I also use as control variables. It also records the currents district of residence as well as the district of birth, that I use to match the information on violence.

In particular, I investigate the degree of trust over the army, the police, parliament, political parties, judiciary system and church. Regarding identity, I investigate whether individuals feel closer to their neighbors (locals), their ethnic/race peers or their religious groups.

The question about trust lists several government institutions and then inquires the individuals in the household aged 16 or more about how much they trust such institutions. The respondents are offered five alternatives: (a) no trust at all, (b) little trust, (c) just enough, (d) a lot, and (e) do not know. I re-categorize each variable to reflect 1 for responses (c) and (d), and 0 for responses (a) and (b).

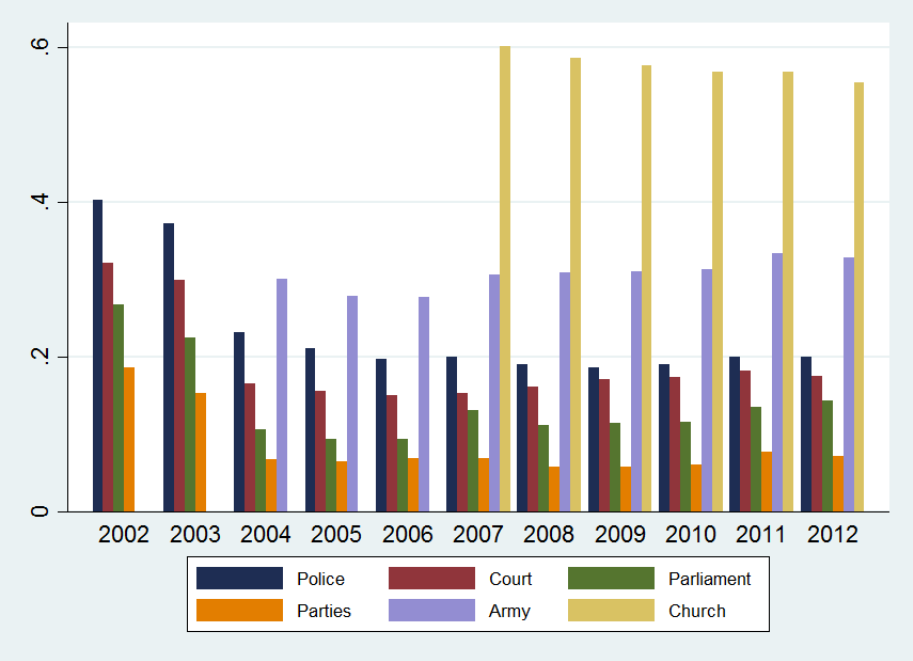
ENAHO, collected information about beliefs since 2002. I cannot, however, use the whole period for the analysis because of changes in the framing of the questions and categorization of the alternatives. Figure 1.6 plots the annual averages for the six variables on trust I use in the empirical analysis. Only four institutions, the police, the parliament, the judiciary system (national court) and political parties were considered for the questionnaire since 2002.

The wording of the alternatives to measure the degree of trust changed through the years, which creates problems for the harmonization of the answers. For the years 2002 and 2003, to the question *do you trust institutions like... ?* the alternatives for each institution were ranked from positive to negative: (a) a lot, (b) more or less, (c) little and (d) not at all. Which in principle looks just like an inverted order as the alternatives for the same question I utilize (period 2007-2012) and explained above: (a) no trust at all, (b) little trust, (c) just enough, (d) a lot, and (e) do not know. However, as Bertrand and Mullainathan (2001) suggest, these

differences are non trivial and actually influence individual's responses to the same question. This becomes evident in 2004, when the order of the alternatives was reverted and the wording for the alternative more or less was changed to just enough, the degree of trust to the police, for instance drops importantly from 0.37 in 2003 to 0.22 in 2004. This sudden change is explained more likely by the different framing rather than an expected change in trust toward the police.

Also from 2004 the army was included in the list of institutions. The church was included among the list of institutions in 2007. This is the reason why I end up using 2007-2012 as the reference period, where I observe the same exact list of institutions and wording/ranking of the alternatives. Appendix B.1 explains the derivation of the trust variables from the translated question that applies to the period 2007-2012 and shows the basic stats.

Figure 1.6: Historical Trust



Notes: [1] ENAHO 2002-2012

In general, it is clear that the level of trust towards public institutions in Peru is low. For instance, the proportion of people who trusts political parties is consistently the lowest in all years. The church stand as the most trusted institution in Peru, which coincidentally is not a public institution.

The way I measure identity in this chapter derives from the only question available in

ENAHO related to this dimension[16]. Identity could take many forms. The most common perhaps is the feeling of belonging to a nation (Masella; 2013), but of course it could take the form of belonging to a social class, a neighborhood, or a racial group. The identity question I end up using is designed to address this multidimensionality in some way. The question inquired individuals about the group (or community) they felt most identified with. Although loosely defined, the mutually exclusive alternatives were: (a) your region, province, district or local community, (b) your ethnic group or race, (c) your peasants or indigenous community, (d) your group or religion beliefs, and (e) others.

The first alternative offers a rather vague geographical sense of contentedness which could also refer to the feeling of belonging to a regional body (region). This regional body also offers a sense of authority or institutional arrangement. The main authority in Peru is the president who symbolizes the nation. The country however, also elects regional presidents who are in charge of major infrastructure works in the region. The role of regional president requires some degree of coordination with the province and district majors. I argue that this category symbolizes the sub-national identity, but with a local reference, since the term "local community" is also embedded in this category. Ultimately this category signifies some local figure related to some form of authority. The importance of establishing this category can be understood in the context of a dichotomy between some sort of regional or national sense of belonging against a sense of narrowly defined ethnic identity.[17]

From table 1.4, for instance the first row indicates that districts with a large proportion of land in communal setting (the dimension of alternative c) are associated with large proportions of people speaking *Quechua* as main language. For both census waves (1994 and 2007) the proportion of *Quechua* speakers is larger in districts of large communal settings. Moreover, table 1.5 immediately signals the strong correlation between native districts (with a large components of native individuals) and the proportion of individuals speaking *Quechua* or other native language. This correlation is not trivial, since one would suspect that people with native origins may speak Spanish as a way to climb in the social ladder. What this table

[16]The transcription of the question and the derivation are explained in appendix B.2

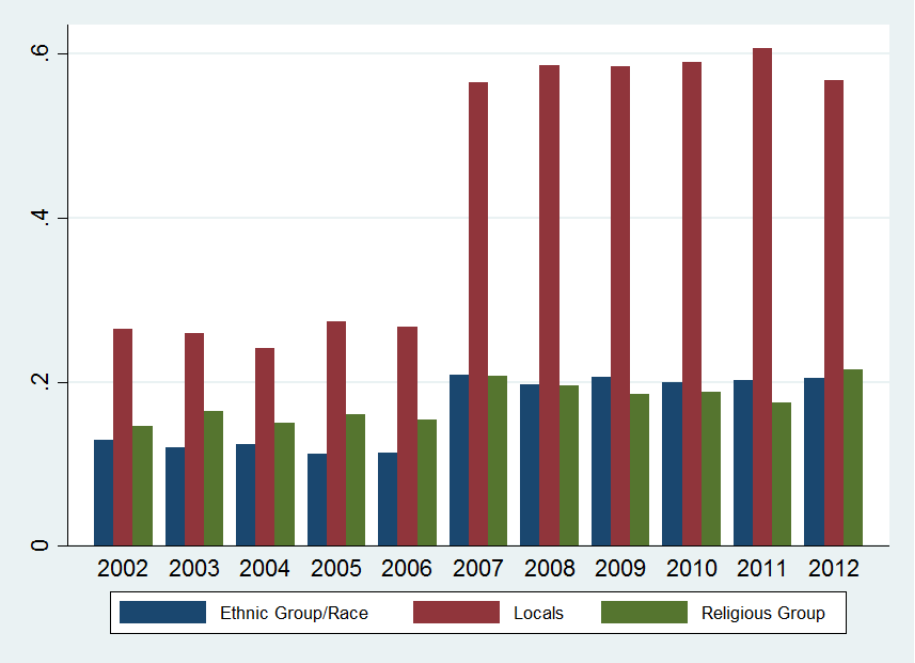
[17]As for instance defined by Miguel (2004). Eifert et al. (2010) find that closer to an election, individuals in Africa feel more identified with their ethnic roots rather than to occupational group or social class. The implications of assuming one type of identity over social preferences have been studied by Chen and Li (2009).

highlight is the prevalence of native languages as main language among indigenous population. This is again evident from tables 1.6 and 1.7 that use individual level data to provide a first glimpse of the association between categories b and c of the question about identity.

Finally, category d refers to any sort of religious symbol. The main religion in Peru is Catholic, however in recent years Evangelist religion has gained support. Unfortunately from the wording of the alternative I cannot distinguish between types of religions or whether it refers to people taking a pro-active attitude or simply a sense of belonging to a group of religious people.

In general, the question about identity also started in 2002, but as in the previous case, the framing and number of alternatives changed in 2007. The most noticeable change is the definition of locals, which included all Peruvians in the 2002-2006 version and therefore would refer to a feeling of belonging to a nation rather than something sub-national. For the period 2007-2012 it covers up to citizens of the region as the largest geographical unit, which was described in the paragraph above. I decided to keep the same period for the three variables on identity: local (or sub-national), ethnic or race, and religion. Historically the proportion if individuals who felt more identified with their locals has been the highest (see figure 1.7).

Figure 1.7: Historical Identity



Notes: [1] ENAHO 2002-2012

Table 1.6: Summary Stats: Individual Beliefs by Native Speaker Condition

	Non-Native	Native	Total
Army	35.25	23.32	32.61
	(47.78)	(42.29)	(46.88)
Police	21.12	15.74	19.93
	(40.82)	(36.42)	(39.95)
Jury	18.25	13.39	17.17
	(38.62)	(34.05)	(37.71)
Parliament	13.58	10.16	12.82
	(34.25)	(30.21)	(33.43)
Parties	6.844	6.409	6.747
	(25.25)	(24.49)	(25.08)
Church	60.01	50.46	57.90
	(48.99)	(50.00)	(49.37)
Ethnic Group/Race	13.26	33.90	17.83
	(33.92)	(47.34)	(38.28)
Locals	64.59	49.64	61.28
	(47.83)	(50.00)	(48.71)
Religion	20.54	15.46	19.42
	(40.40)	(36.16)	(39.56)
Observations	104635		

Notes: [1] Data source: ENAHO 2007-2012. [2] Unweighted averages. Standard deviations shown in parenthesis. [3] An individual is native if their mother tongue is of any native origin: Quechua, Aymara or other

Table 1.6 presents the summary statistics for the trust and identity variables broken down by the native origin of the individual. In general, Peruvian individuals exhibit low levels of trust on the institutions included in the analysis. The national proportion of individuals who trust the army is around 32.61%, with native individuals showing a lower level of trust than non-native. The police is even less trusted: the national average is 19.93%, again, native individuals show a smaller trust on police, 15.74%, compared to a higher trust from non-native individuals, 21.12%. A potential explanation for this difference is the crime rate between urban and rural regions, with the first being more populated than the second ones. The judiciary system is yet less trusted: the national average reaches 17.17%, again with native individuals showing smaller levels of trust than non-native individuals: 13.39% vs 18.25%. The proportion of individuals who trust the Parliament is just 12.82%, again, when split by the native status, native individuals show a smaller level of trust, 10.16%, than non-native individuals 13.58%. At the bottom of the preferences, the political parties can only make 6.74% of the population trust them. This time there is little difference depending on the native status of the individual. Perhaps related to this general distrust to government or official institutions, the church seems to harvest most the trust of Peruvian individuals. The proportion of Peruvians who trust the church is about 57.90%. This time, native individuals show a smaller level of trust than the non-native: 50.46% against 60.01%.

Regarding identity, most of Peruvians identify themselves with their local neighbors: 61.28%. This sense of identity is higher for non-native individuals, 64.59%, than for non-native, 49.64%. The national proportion of Peruvians who identify with their ethnic group or race is 17.83%. Not surprisingly, this sense of connection is higher for native individuals, who in their majority populate the highlands: 33.9%. Just 13.26% of the non-native individuals declare that they identify with their ethnic group or race. Identification with religious groups is also important among Peruvian individuals. The national proportion of individuals who feel identified with any religious group is 19.42%. Non-native individuals express a higher degree of identification with religion, 20.58%, than native individuals, 15.46%.

If instead of the native origin of the individual, the classification uses the proportion of agricultural land under the tenancy of peasants or indigenous communities the differences remain. In this case, a district is classified as communal if the share of agricultural land

owned by peasant or indigenous communities is higher than 0.5. Table 1.7 shows the differences. As in the native classification, individuals living in communal districts trust less the government institutions, and identifies themselves more with their ethnic or racial group.

Table 1.7: Summary Stats: Individual Beliefs by Communal District Status

	Non-Communal	Communal	Total
Army	34.52	28.61	32.60
	(47.54)	(45.19)	(46.87)
Police	20.42	18.85	19.91
	(40.31)	(39.11)	(39.93)
Jury	17.78	15.88	17.17
	(38.24)	(36.55)	(37.71)
Parliament	13.22	11.99	12.82
	(33.87)	(32.48)	(33.43)
Parties	6.654	6.940	6.746
	(24.92)	(25.41)	(25.08)
Church	58.94	55.75	57.91
	(49.19)	(49.67)	(49.37)
Ethnic Group/Race	14.17	25.43	17.82
	(34.87)	(43.55)	(38.27)
Locals	63.59	56.56	61.31
	(48.12)	(49.57)	(48.70)
Religion	20.58	16.97	19.41
	(40.43)	(37.53)	(39.55)
Observations	104425		

Notes: [1] Data source: ENAHO 2007-2012. [2] Unweighted averages. Standard deviations shown in parenthesis. [3] A district is referred as communal if the share of agricultural land under communal tenancy is larger than 0.5

ENAHO also collects information on the individual characteristics I include as controls: (i) the number of years of schooling, (ii) gender (male = 1), (iii) a dummy variable to indicate whether the individual learned any native language as mother tongue, and a (iv) a dummy to indicate that the individual resides in the coastal region of Peru. The empirical equation also includes age fixed effects.

Finally, local mobility of individuals from the rural highlands, where the violence was more frequent, is relatively low. The proportion of individuals living in a district different from the district of birth is 31%. This, however may imply that individuals relocate to somewhere nearby. The percentage of individuals living in a province (a higher administrative unit) different than the province of birth is 23%. While the migration rate at the regional level is 15%. Adhvaryu and Fenske (2014) find that controlling for migration in different

ways does not alter the results. I would expect that if conflict had an effect on migration, it was temporary, and that the district of birth fixed effect in the empirical section should account for different migration rate by districts affected by the violence.

1.4 Empirical Framework

The main specification, hence, takes the following form:

$$y_{irt} = \alpha + \beta * killed_{i,ro} + \gamma_{age} + \gamma_p(t) + \theta_r + \theta_{ro} + \psi_t + X_{it} + \varepsilon_{irt} \tag{1.2}$$

Where y_{irt} is the belief. γ_{age} is age fixed effect, θ_{ro} is district of birth fixed effect, while θ_r is current district fixed effect. ψ_t is year fixed effect, $\gamma_p(t)$ is a set of current province-specific trends. X_i is a vector of individual characteristics: the years of schooling, gender (male), a dummy for individuals whose mother tongue is either *quechua* or *aymara*, a dummy for individuals living in the coast region. The period of analysis is 2007 to 2012. ε_{irt} is the error term, clustered at the district of origin level. After controlling for common fixed effects at the district of origin, the district of residence, year fixed effects, province trends and age (cohort) fixed effects, this methodology uses the exposure to violence tat is not explained by these set of fixed effects. β captures such effect: the exposure to violence, as defined in 1.1. I evaluate the effect of total violence (both army and terrorist) and also the effect of army and terrorist violence separately.

To differentiate by the identity of the perpetrator, I used the re-definition of the variable $killed_{i,ro}$ where I count impressionable years affected for army violence only and terrorist violence only. I present the results in separate regressions for each violence definition.

A second specification interacts the number of impressionable years, $killed_{i,ro}$ with the ethnic origins of the individual. where the ethnic origin is measured in the three different ways explained in section 1.2.2: the share of communal land, the native origin, or the *comunero* status:

$$
\begin{aligned}
y_{irt} &= \alpha + \delta_1 * killed_{i,ro} + \delta_2 * killed_{i,ro} * ethnic_{i,ro} \\
&\quad + \gamma_{age} + \gamma_p(t) + \theta_r + \theta_{ro} + \psi_t + X_{it} + \varepsilon_{irt} \qquad (1.3)
\end{aligned}
$$

The coefficients δ_2 then captures the effect of violence during the impressionable years of the individuals interacted with the measure of ethnic origins. For this specification I also differentiate by the identity of the perpetrator: the government or the terrorists.

Before proceeding to the results, it is worth discussing potential selection problems. The identification strategy aims to address problems of selection through the use of the spatial progression of the conflict. The first concern relates to selection. It would be an estimation problem if conflict was originally targeted to districts with a characteristic degree of trust and identity. If this was the case, the subsequent effect I attach to conflict would be instead the simple selective nature of the conflict.

Precisely ruling this possibility is difficult due to data constraints, but I will argue that conflict was not targeted to districts with a characteristic degree of trust toward institutions and identity with respect to some groups.

I do not observe trust and identity before 2007, therefore I cannot directly test whether by the onset of the conflict districts were targeted according to their degree of trust or identity. However, I will argue that if violence had a determined geographical target, it should have focused on it through the whole duration of the conflict. Instead, figure 1.8 shows that the spatial progression of the conflict shifted its intensity from the South-Central (blue bars) region to the capital Lima (yellow bars).

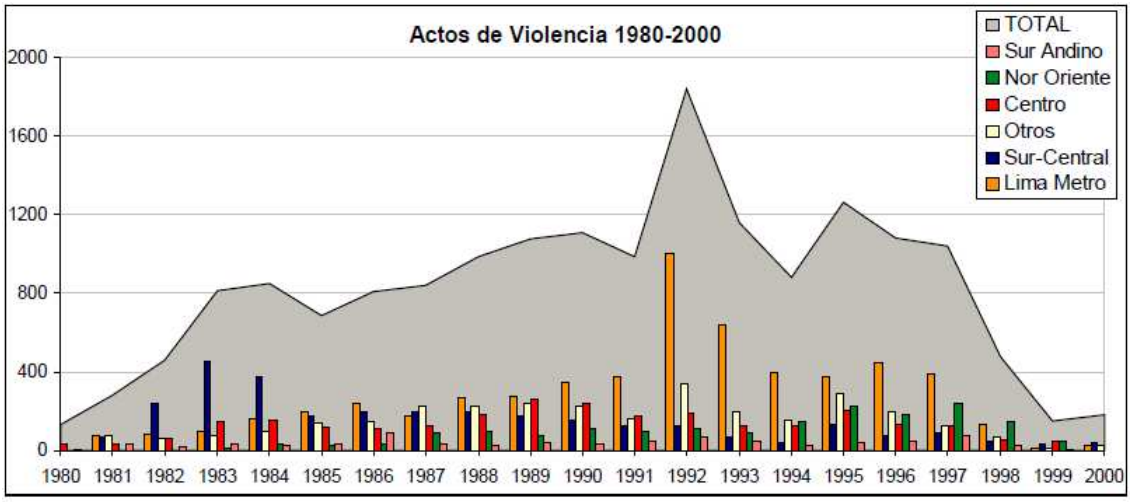
The conflict started in Ayacucho, which is a city located in the South-Central part in the country, and that is reflected in the number of events focused in that region by the early years. Ayacucho is allegedly a region with a high proportion of native speakers (88% in 1981), which would identify more with their ethnic groups. At the same time, is a predominantly rural region. However, by early 90s the conflict shifted to Lima, the capital, center of financial power, with no agricultural production and populated by Spanish speakers (85%) who would probably not identify themselves with any native ethnic group.

This idea is well described by the CVR (2003) in chapter that describes the spatial progression of the conflict. According to the description in that chapter, the conflict initially targeted the South-Central region (where Ayacucho is) and looted government infrastructure like municipal buildings, police stations, electricity towers, dams, among others. Therefore, even if the initial geographical target of the violent action was a region dominated by indigenous population, the guerrilla explicitly sought to destroy national government institutions and infrastructure.

This becomes more evident after the spatial progression of the conflict, when the infrastructure targets were the same, but the attacks were perpetrated in different regions.

Finally, the empirical strategy precisely exploits this spatial progression. In concrete, it compares two individuals of similar characteristics who were exposed to violence during the impressionable years with different degrees. For instance, it would compare two individuals currently living in the north of the country (where the conflict was rare) but with different exposure to violence during their formative period. Individual A who lived in Ayacucho when she was 18 years old, and individual B who lived in another southern region also when she was 18 years old, less exposed to conflict. After controlling for current district fixed effects, district of origin fixed effects, current province trends, and the remaining controls specified in equations 1.2 and 1.3, I expect that the differential exposure to conflict when both individuals were 18 years old explains some differences in their attitudes regarding trust and identity.

Figure 1.8: Spatial Progression of Armed Conflict



Notes: [1] Data source: CVR (2003). [2] Vertical axis measures the number of violent events in the year. Horizontal axis plots the regional frequency by year.

1.5 Results and Discussion

1.5.1 The effect of conflict on trust and identity

The inclusion of many controls and fixed effects leaves little variation to exploit for the estimation of the results. However, I prefer to follow the literature in this regard and leave these controls and fixed effects and interpret the results based on the few results that are statistically significant, which is the best I can do in this context.

Table 1.8 presents the results of estimating equation 1.2 for the group of variables on trust. Panel (a) evaluates the baseline specification with the number of impressionable years as explanatory variable. Panels (b)and (c) use an alternative measure for exposure to violence. Panel (b) uses exposure to violence measured as dummy variable: at least one impressionable year affected by violence. Panel (c) attempts to measure the intensity of the conflict that the individual was exposed to during their impressionable years: the specification here uses the number of victims in the district during the impressionable years of the individual scaled by the district population in 1993 (in thousands). I report results for overall violence exposure (columns labeled *All*), and exposure to army violence (columns labeled *A*) and terrorist violence (label *T*) alone.

Table 1.8: Trust

	Army			Police			Jury			Parliament			Parties			Church		
	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
a. Number of Impressionable Years																		
Killed	-0.19	-0.39	-0.24	-0.25*	-0.62***	-0.32**	-0.28**	-0.44**	-0.33**	-0.20*	-0.56***	-0.23*	-0.05	-0.09	-0.05	-0.19	-0.48	-0.37*
	(0.16)	(0.26)	(0.20)	(0.14)	(0.21)	(0.16)	(0.12)	(0.21)	(0.14)	(0.11)	(0.18)	(0.13)	(0.09)	(0.15)	(0.10)	(0.18)	(0.32)	(0.19)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
b. At Least One Impressionable Year																		
Killed (0/1)	-0.42	-0.36	-0.69	0.10	-0.53	-0.18	-0.77*	-0.65	-0.47	0.19	-0.52	0.23	-0.08	0.13	-0.17	0.00	0.56	-0.09
	(0.61)	(0.82)	(0.64)	(0.47)	(0.60)	(0.51)	(0.46)	(0.64)	(0.45)	(0.41)	(0.65)	(0.46)	(0.34)	(0.51)	(0.33)	(0.64)	(0.81)	(0.71)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
c. Intensity																		
Killed	0.26	0.62**	-4.52	-0.03	0.26	-4.21**	0.46	0.77**	-3.22**	-0.35	-0.15	-3.03	0.41	0.61**	-1.66	-0.31	-0.29	-1.09
	(0.40)	(0.31)	(3.28)	(0.74)	(0.96)	(1.69)	(0.34)	(0.32)	(1.50)	(0.48)	(0.57)	(1.92)	(0.41)	(0.30)	(1.12)	(0.42)	(0.43)	(2.28)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115431	115431	115431	120527	120527	120527	114357	114357	114357	111434	111434	111434	116029	116029	116029	124567	124567	124567

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

Results in table 1.8 indicate that there is an statistically significant effect when violence is measured as the sum of impressionable years exposed to violence, as reported in panel (a). The effect is negative and statistically significant for trust towards the police, the judiciary institution and the parliament. The estimated coefficients are very low, however, when compared to the dependent variable means acquire economical meaning. All of them are less than 1%. The average number of impressionable years exposed to violence was 0.23, with an standard deviation of 0.97.

Focusing on panel (a), a one standard deviation increase in the number of impressionable years exposed to overall violence reduces the probability to trust the police by 0.24 (0.25x0.97) percentage points. Compared to the average trust in the police, 19.93, the estimated effect is about 1% of the mean.

The estimated effect becomes slightly stronger when the estimation isolates the violence exerted by the army: a one standard deviation increase in the number of impressionable years exposed to government violence reduces the probability to trust the police by 0.60 (0.62x0.97) percentage points (3% of the average trust). Terrorist violence alone explains a reduction in trusts towards the police by 0.31 (0.32x0.97) percentage points (1.5% of the average trust in police).

The effect of violence on trusts over the judiciary institution and the parliament exhibits the same pattern: a negative overall effect where the violence exerted by the government galvanizes a stronger effect. In the case of the trust on church, only terrorist violence has an effect that is statistically significant at 10%. Overall violence reduces trust by 0.43 (0.44x0.97) percentage points, which compared to the average trust on the judicial system (17.2%), represents a non-negligible 2.5%.

In general, the average mean of the dependent variables is small, which grants the estimated coefficients some economic importance despite of their small estimated effect

Panel (b) shows a similar patter, however only the overall effect of violence on trust towards the judiciary institution is statistically significant at 10%. Results in panel (c) confirm the negative effect of terrorist violence over trust towards the police and the judiciary institution. But the intensity of army violence is associated with a positive effect on trust towards the army, the judiciary system and the political parties.

To understand these results, it is worth referring to the comments of the CVR on the consequences of the violence in Peru. According to their account, the judiciary was among the less trusted institutions during the conflict, which is explained by the shallow reach of government institutions in general in the regions of the conflict. It is remarkable, at the same time, that there is no effect on trust towards the army in any of the specifications.

Regarding identity, table 1.9 presents a similar set of results. The main effect to highlight here is the negative effect of violence on the probability of identification with the group of locals and the positive effect on the probability of identification with the religious groups. More importantly, in both cases the effect when statistically significant is carried by terrorist violence. There is also a consistent positive effect of violence on identification with the ethnic and race group, however none of the estimated coefficients is statistically significant at any reasonable level. However, as in the case of the effect of violence over trust, the estimated effects are relatively small.

Panel (a) indicates that overall violence reduced the probability that an individual identifies with the local population, however this result is only statistically significant at 10%. In concrete, a one standard deviation increase in the number of impressionable years affected by overall violence reduces the probability of local identification by 0.37 percentage points (0.38x0.97). The same violence has the opposite effect on the probability of religious identification: a one standard deviation increase in the number of impressionable years affected by overall violence increases the probability that a random individual identifies him self or herself with a religious group by 0.30 (0.31x0.97) percentage points. In this case, the effect is driven by terrorist violence. The estimated coefficient for army violence is smaller in magnitude and not statistically significant.

Results in panel (b) show the similar pattern. Having experienced at least one impressionable year of overall violence results in an increase of the probability of local identification by 2 percentage points. Compared to the average degree of identification with locals (17%), this effect is about 11%. As in the case with the estimated coefficients on trust, the small averages for the dependent variables grants economic significance to the estimated effects.

When the estimation looks for an heterogeneous effect depending on the identity of the perpetrator, terrorist violence stands out over army violence: the coefficient is not only statis-

Table 1.9: Identity

	Local/Neighbour			Ethnic/Race			Religion		
	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
a. Number of Impressionable Years									
Killed	-0.38*	-0.26	-0.34	0.12	0.18	0.11	0.31**	0.11	0.30**
	(0.22)	(0.47)	(0.27)	(0.16)	(0.29)	(0.19)	(0.13)	(0.27)	(0.15)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
b. At Least One Impressionable Year									
Killed (0/1)	-2.00***	-1.06	-1.80***	0.74	0.72	0.34	1.32***	0.12	1.49***
	(0.62)	(0.87)	(0.67)	(0.47)	(0.70)	(0.53)	(0.50)	(0.66)	(0.50)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
c. Intensity									
Killed	-0.47*	-0.54**	0.37	0.71*	0.88***	-1.35	0.22	0.07	2.66
	(0.28)	(0.21)	(2.81)	(0.38)	(0.27)	(3.83)	(0.27)	(0.20)	(2.24)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126194	126194	126194	126194	126194	126194	126194	126194	126194

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

tically significant but higher than the coefficient for army violence (which is not statistically significant). Al least one year of terrorist violence during the impressionable years reduces the probability of feeling identified with the local population by 1.8 percentage points.

In the case of identification with religious groups the effect is the opposite but with the same pattern: overall violence during the impressionable years increases the probability that the individual identifies with religious groups by 1.32 percentage points. Such effect is explained by terrorist violence: at least one year of terrorist violence during the impressionable years increases the probability of finding identification with religious groups by 1.49 percentage points.

Panel (c) offers coefficients with the same sign as in the previous panels. The intensity of army violence had a negative impact on the degree of identification with the local population: -0.54. A one standard deviation increase in the number of victims by population

(215.5[18]) reduces the probability if identification with locals by 116.1%. This result is quite extreme, and may be the result of including extreme values in the ratio of victims over population: the average army victims over population is 3.5, while the median is zero and the 95th percentile is also zero. Therefore, although the estimations using the intensity are illustrative, the estimation of the coefficients may be subjected to extreme values[19].

How to understand the identity results? The CVR again offers guidance in this regard. In the chapter about the psychological consequences of the violence[20]. According to the accounts of CVR (2003), religious groups were particularly supportive during the conflict period in many regions in the highlands. This comes as a natural consequence of the absence of government institutions in those regions, where ultimately other types of informal institutions supplanted the government's role. This result will be clearer in section 1.5.2.

Conflict During Early Childhood

In their account, Bauer et al. (2016) suggest that the effect of violence on cooperation behavior varies little depending on the age that the individual was exposed to the conflict. In a mostly related work, Adhvaryu and Fenske (2014) investigate the formation of beliefs in African individuals exposed to conflict during their early childhood and finds a small positive effect for generalized trust.

In order to shed more light on the results of this chapter, this section evaluates whether the choice of a different life period offers different results. In particular, I focus on the period when the individual was aged between 0 and 15 years. Violence during that period of life could have different effects by means of different factors. In this section I do not account for the channels through which violence at early stages may affect trust and identity beliefs. However, one possible channel is education. As found by León (2012), individuals who experienced the violence period at early childhood accumulate 0.31 less years of education as adults.

Empirically, I re-estimate equation 1.2 but using the number of early years (between 0

[18]The average and standard deviation for the ratio of army victims/population are 3.5 and 215.5, respectively.
[19]Also, this is a linear probability model, in which case the predicted probability could reach extreme values at both extremes.
[20]CVR (2003), Part III, chapter 1.

and 15) affected by violence as explanatory variable.

Table 1.10 presents the results for trust. Experiencing violence during early childhood has a different effect on trust. First, the effect on the judiciary institution and the political parties is gone. Second, the effect on trust over the police changes sign and becomes positive. Third, there is now a positive effect of violence on trust over the army. Finally, the church is less trusted by individuals exposed to violence during their impressionable years.

Table 1.10: Trust: Early Childhood

	Army			Police			Jury			Parliament			Parties			Church		
	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
a. Number of Early Years																		
Killed	0.33**	0.55	0.37*	0.35***	0.55***	0.46***	0.16	0.17	0.19	0.06	0.28	0.10	0.11	0.18	0.15	-0.32**	-0.56*	-0.34*
	(0.17)	(0.38)	(0.19)	(0.09)	(0.20)	(0.11)	(0.14)	(0.26)	(0.18)	(0.10)	(0.21)	(0.12)	(0.08)	(0.15)	(0.09)	(0.15)	(0.30)	(0.19)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
b. At Least One Early Year																		
Killed (0/1)	0.94	1.01	1.20*	0.90	0.80	1.21*	-0.34	-0.01	-0.04	-0.27	-0.10	0.13	0.07	0.32	0.33	-0.78	-2.13***	-0.81
	(0.73)	(1.11)	(0.73)	(0.59)	(0.83)	(0.64)	(0.59)	(0.98)	(0.65)	(0.53)	(0.76)	(0.56)	(0.33)	(0.43)	(0.33)	(0.65)	(0.82)	(0.70)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
c. Intensity																		
Killed	0.61	0.52	2.60	1.32***	1.26***	3.62**	-0.32	-0.56	2.92	-0.25	-0.42	1.99	-0.09	-0.32*	3.10**	-0.74*	-0.37**	-6.27**
	(0.47)	(0.47)	(1.82)	(0.18)	(0.21)	(1.75)	(0.59)	(0.73)	(2.51)	(0.35)	(0.28)	(1.62)	(0.21)	(0.17)	(1.55)	(0.44)	(0.15)	(2.76)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115431	115431	115431	120527	120527	120527	114357	114357	114357	111434	111434	111434	116029	116029	116029	124567	124567	124567

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of early years (between 0 and 15) affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

These positive effects open a different discussion with respect to current result. However, the results at early age they are much in line with the literature has found on the effect of violence on cooperative behavior. As mentioned earlier, Bauer et al. (2016) in their meta-analysis find that exposure to violence leads to an increase in pro-social behavior[21]. However, within the set of result they present, the effect on trust on institutions is not statistically significant. The result of this section not only offer a validation to their findings applied to a different context, but also provide additional evidence that the experience of violence at different stages in life may have different results during adult life.

In line with the above description, the results on identity are also different. Table 1.11 suggest that exposure to violence during early childhood has the opposite effect on the identification measures as the ones found in table 1.9.

Table 1.11: Identity: Early Childhood

	Local/Neighbour			Ethnic/Race			Religion		
	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
a. Number of Early Years									
Killed	0.06	0.32	-0.03	-0.18	-0.36	-0.17	0.11	0.05	0.21
	(0.17)	(0.36)	(0.19)	(0.16)	(0.32)	(0.19)	(0.14)	(0.24)	(0.15)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
b. At Least One Early Year									
Killed (0/1)	-0.04	1.39	-0.66	-0.93	-1.65**	-0.81	0.91	0.14	1.35**
	(0.61)	(1.02)	(0.65)	(0.57)	(0.84)	(0.61)	(0.56)	(0.83)	(0.57)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
c. Intensity									
Killed	1.46***	1.70***	0.78	-0.66	-0.57	-2.72	-0.48*	-0.72***	1.83
	(0.33)	(0.32)	(2.14)	(0.46)	(0.44)	(2.72)	(0.29)	(0.15)	(1.58)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126194	126194	126194	126194	126194	126194	126194	126194	126194

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of early years (between 0 and 15) affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

[21] Adhvaryu and Fenske (2014) also provide evidence in this direction, and as in the present case, the general effect o the measure they have for trust, is very small.

Proliferation of Fixed Effects

The estimation strategy is demanding in the number of fixed effects it requires. The possibility that the estimated coefficients are to small may be related to this, so in this section I provide some estimates of simple correlations to understand the nature of the link between conflict and beliefs. In concrete, I re-run the empirical equation 1.2 excluding the dummies for the district of origin or the district of residence, neither the province-specific trends. Tables 1.12 and 1.13 show the results. The main conclusion regarding trust is that the estimated coefficients are slightly larger and all of them with the negative sign. In the case of identity, there are no statistically significant effect, and the coefficients estimated in panel (c) are practically zero. These results suggest that controlling for geographical fixed effects at residency and birth, as well as province trends, reduces the effect of violence over trust.

Table 1.12: Trust: No Fixed Effects

	Army			Police			Jury			Parliament			Parties			Church		
	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
a. Number of Impressionable Years																		
Killed	-1.48***	-2.66***	-1.70***	-1.09***	-1.93***	-1.29***	-0.93***	-1.44***	-1.06***	-0.71***	-1.21***	-0.78***	-0.55***	-0.88***	-0.67***	-0.33	-0.66	-0.51
	(0.27)	(0.41)	(0.34)	(0.17)	(0.24)	(0.21)	(0.17)	(0.27)	(0.20)	(0.15)	(0.28)	(0.19)	(0.13)	(0.16)	(0.15)	(0.37)	(0.70)	(0.46)
R^2	0.05	0.05	0.05	0.01	0.01	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03
Observations	54734	54734	54734	57978	57978	57978	54332	54332	54332	51655	51655	51655	54786	54786	54786	61140	61140	61140
b. At Least One Impressionable Year																		
Killed (0/1)	-4.66***	-5.44***	-5.47***	-3.02***	-4.03***	-3.68***	-3.33***	-3.38***	-3.36***	-1.88***	-3.20***	-2.12***	-1.52***	-1.74**	-1.88***	-0.48	-0.91	-1.08
	(1.07)	(1.44)	(1.21)	(0.75)	(0.94)	(0.82)	(0.77)	(0.90)	(0.80)	(0.63)	(0.77)	(0.68)	(0.51)	(0.70)	(0.51)	(1.25)	(1.70)	(1.28)
R^2	0.05	0.05	0.05	0.01	0.01	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03
Observations	54734	54734	54734	57978	57978	57978	54332	54332	54332	51655	51655	51655	54786	54786	54786	61140	61140	61140
c. Intensity																		
Killed	-0.35	0.19	-7.44**	-0.72	-0.30	-6.81***	-0.09	0.22	-4.14***	-0.71*	-0.40	-4.93***	-0.09	0.23	-3.68***	-1.30***	-1.27***	-3.97
	(0.58)	(0.17)	(3.56)	(0.74)	(0.95)	(1.14)	(0.28)	(0.35)	(1.08)	(0.41)	(0.44)	(1.67)	(0.46)	(0.25)	(0.88)	(0.33)	(0.26)	(2.64)
R^2	0.03	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02
Observations	115431	115431	115431	120527	120527	120527	114357	114357	114357	111434	111434	111434	116029	116029	116029	124567	124567	124567

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include as controls: dummies for years and age, years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

Table 1.13: Identity: No Fixed Effects

	Local/Neighbour			Ethnic/Race			Religion		
	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
a. Number of Impressionable Years									
Killed	0.32	0.41	0.26	-0.56*	-0.46	-0.55*	0.27	0.09	0.32
	(0.25)	(0.38)	(0.29)	(0.31)	(0.40)	(0.33)	(0.20)	(0.30)	(0.22)
R^2	0.09	0.09	0.09	0.13	0.13	0.13	0.04	0.04	0.04
Observations	62198	62198	62198	62198	62198	62198	62198	62198	62198
b. At Least One Impressionable Year									
Killed (0/1)	1.84*	1.14	2.00*	-2.78**	-1.47	-2.66**	0.87	0.42	0.62
	(1.04)	(1.49)	(1.13)	(1.26)	(1.74)	(1.27)	(0.83)	(1.02)	(0.78)
R^2	0.09	0.09	0.09	0.13	0.13	0.13	0.04	0.04	0.04
Observations	62198	62198	62198	62198	62198	62198	62198	62198	62198
c. Intensity									
Killed	-0.00***	-0.00***	-0.01**	0.00	0.00	0.00	0.00***	0.00***	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.06	0.06	0.06	0.09	0.09	0.09	0.04	0.04	0.04
Observations	126194	126194	126194	126194	126194	126194	126194	126194	126194

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include as controls: dummies for years and age, years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

1.5.2 The effect of conflict and the communal origins of the individual

In this section I empirically test whether the exposure to violence during the impressionable years had a different result depending on the indigenous origin of the individuals and I provide an answer to the question whether individuals with some degree of connection to the indigenous dimension in Peru have a different reaction to the violence. In a way, this provides also an answer to the question about whether people who belong to groups persistently poor and historically neglected develop a higher or lower trust over institutions, or whether they have a different sense of connection to their peers.

I measure the degree of connection to the indigenous dimension in three ways, as explained in section 1.2.2: (i) with the share of agricultural land in hands of peasant or indigenous groups or associations, (ii) with the *comunero* characteristic of the individual, and (iii) with the native characteristic of the individual. These three ways of measurement are represented by panels (a), (b), and (c), respectively, in tables 1.14, 1.15, 1.16, and 1.17.

Tables 1.14 and 1.15 use the empirical specification for equation 1.3 that considers exposition to violence measured as the number of impressionable years under violence.

Regarding trust (table 1.14), the indigenous interaction suggests that individuals with some degree of connection to the indigenous dimension and who were exposed to violence during their impressionable years report a smaller probability to trust political parties. This result is evident in columns 13, 14 and 15 for panels (a) and (c). For instance, taking panel (a) for the interpretation, and column 13, one standard deviation increase in the number of impressionable years exposed to overall violence led to a reduction in the probability of trusting the political parties by 0.26 ([0.16 - 0.43]*0.97) percentage points for individuals born in a district where more than half of the agricultural land was held by peasant or indigenous communities. Although small, this result represents 4% of the average trust on political parties (0.26% out of 6.7%). Panel (c) provides suggestive evidence that being a *comunero* increases the magnitude of the negative effect of violence during the impressionable years over the individual trust over the parliament. In particular, the experience of at least one impressionable year affected by army violence. In this case, the estimated effect or army violence considering being a *comunero* reaches -1.58 (in column 11, panel (c): -1.58 = -0.48 -

1.07).

Table 1.14: Trust Interacted Number of Years

	Army			Police			Jury			Parliament			Parties			Church		
	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
a. Share																		
Killed	-0.10	0.10	-0.17	-0.35*	-0.75***	-0.44**	-0.47**	-0.61*	-0.52***	-0.08	-0.22	-0.05	0.16	0.35*	0.23*	-0.30	-0.63	-0.54**
	(0.19)	(0.30)	(0.23)	(0.19)	(0.27)	(0.22)	(0.18)	(0.34)	(0.20)	(0.13)	(0.26)	(0.17)	(0.11)	(0.21)	(0.12)	(0.23)	(0.46)	(0.24)
Killed x Share	-0.19	-0.96**	-0.14	0.21	0.25	0.24	0.38*	0.34	0.38	-0.26	-0.67**	-0.37*	-0.43***	-0.86***	-0.57***	0.21	0.27	0.33
	(0.29)	(0.45)	(0.35)	(0.24)	(0.38)	(0.27)	(0.22)	(0.39)	(0.25)	(0.17)	(0.30)	(0.21)	(0.13)	(0.24)	(0.15)	(0.33)	(0.61)	(0.36)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
b. *Comunero*																		
Killed	-0.19	-0.37	-0.24	-0.24*	-0.58***	-0.31*	-0.28**	-0.38*	-0.32**	-0.17*	-0.48***	-0.18	-0.03	-0.05	-0.04	-0.23	-0.49	-0.41**
	(0.17)	(0.29)	(0.21)	(0.14)	(0.21)	(0.17)	(0.13)	(0.23)	(0.15)	(0.11)	(0.17)	(0.13)	(0.09)	(0.15)	(0.10)	(0.19)	(0.34)	(0.20)
Killed x Comunero	-0.06	-0.29	-0.04	-0.10	-0.43	-0.14	-0.03	-0.77	-0.14	-0.37	-1.07***	-0.57**	-0.19	-0.42	-0.20	0.45	0.06	0.45
	(0.35)	(0.60)	(0.42)	(0.36)	(0.56)	(0.41)	(0.31)	(0.49)	(0.33)	(0.23)	(0.34)	(0.25)	(0.16)	(0.30)	(0.18)	(0.48)	(0.99)	(0.52)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
c. Native																		
Killed	-0.31*	-0.57**	-0.37	-0.29	-0.77***	-0.35	-0.32**	-0.36	-0.37*	-0.10	-0.43*	-0.09	0.11	0.14	0.16	-0.22	-0.62	-0.49**
	(0.18)	(0.27)	(0.23)	(0.18)	(0.22)	(0.23)	(0.16)	(0.29)	(0.19)	(0.13)	(0.24)	(0.16)	(0.11)	(0.19)	(0.11)	(0.21)	(0.42)	(0.23)
Killed x Native	0.29	0.48	0.28	0.10	0.39	0.06	0.10	-0.21	0.08	-0.26	-0.36	-0.33*	-0.39***	-0.62***	-0.49***	0.07	0.34	0.26
	(0.25)	(0.45)	(0.29)	(0.26)	(0.53)	(0.30)	(0.19)	(0.34)	(0.21)	(0.16)	(0.31)	(0.19)	(0.14)	(0.23)	(0.16)	(0.33)	(0.53)	(0.37)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

However, the most interesting results emerge for the identity variables. Table 1.15 indicates that violence had a statistically significant effects on the probability of identification with locals and ethnic groups that oppose each other. This result appears in panel (b), which measures the indigenous dimension with the individual ownership of agricultural land in a communal setting. One standard deviation increase in the number of impressionable years affected by violence represents a 1.36 ([-0.27 - 1.14]x0.97) percentage points reduction in the probability that an individual who is also a *comunero* identifies himself with his locals. Whereas the same shock represents an increase of 1.72 ([-0.06 + 1.83]x0.97) percentage points in the probability that a *comunero* individual identifies himself with his ethnic or race group. Actually, this effect is even stronger when the estimation considers only the violence exerted by the army: 3.42 percentage points increase, which is 19% of the national average of people who identify themselves with ethnic or racial groups (17.83).

Tables 1.16 and 1.17 repeat the exercise considering the a dummy variable that reflects the exposure to the violence during at least one impressionable year. The results are robust for the variables on identification and stronger in magnitude. What also emerges from this exercise is the fact that violence led by the army had the strongest effect. In concrete, focusing on table 1.17 and taking the results from column 2 and panel (b): having experienced at least one impressionable year under conflict led by the army represented a reduction in the probability that a *comunero* individual feels identified with his local neighbors by 8.49 (-0.48 - -8.01) percentage points. At the same time, army violence measured as dummy, represented an increase of 11.75 (-0.14 + 11.89) percentage points in the probability of a *comunero* identifying himself with his ethnic or race group.

Table 1.15: Identity Interacted Number of Years

	Local/Neighbour			Ethnic/Race			Religion		
	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
a. Share									
Killed	-0.16	0.52	-0.09	-0.01	-0.33	-0.02	0.26	-0.18	0.20
	(0.26)	(0.58)	(0.34)	(0.16)	(0.28)	(0.21)	(0.18)	(0.47)	(0.23)
Killed x Share	-0.44	-1.45*	-0.49	0.25	0.96*	0.26	0.10	0.53	0.18
	(0.41)	(0.85)	(0.52)	(0.30)	(0.50)	(0.37)	(0.23)	(0.54)	(0.27)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
b. *Comunero*									
Killed	-0.27	-0.09	-0.23	-0.06	-0.17	-0.07	0.35***	0.24	0.33**
	(0.22)	(0.46)	(0.27)	(0.15)	(0.25)	(0.18)	(0.13)	(0.29)	(0.16)
Killed x Comunero	-1.14**	-1.85	-1.10*	1.83***	3.70***	1.80**	-0.39	-1.39**	-0.37
	(0.55)	(1.23)	(0.62)	(0.62)	(1.26)	(0.72)	(0.37)	(0.56)	(0.43)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
c. Native									
Killed	-0.25	0.08	-0.19	0.12	-0.00	0.16	0.12	-0.19	0.01
	(0.27)	(0.52)	(0.35)	(0.17)	(0.27)	(0.21)	(0.18)	(0.37)	(0.23)
Killed x Native	-0.29	-0.85	-0.31	0.00	0.44	-0.11	0.41*	0.72	0.60**
	(0.39)	(0.58)	(0.44)	(0.32)	(0.46)	(0.35)	(0.23)	(0.46)	(0.28)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

Table 1.16: Trust Interacted Dummy

	Army			Police			Jury			Parliament			Parties			Church		
	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
a. Share																		
Killed (0/1)	-0.76	0.73	-1.28*	-0.12	-0.62	-0.52	-1.15*	-0.71	-0.57	0.09	0.17	0.42	-0.03	0.79	0.14	-0.23	0.03	-0.24
	(0.74)	(1.01)	(0.75)	(0.50)	(0.64)	(0.56)	(0.61)	(0.83)	(0.57)	(0.46)	(0.68)	(0.54)	(0.47)	(0.81)	(0.47)	(0.77)	(1.07)	(0.95)
Killed (0/1) x Share	0.82	-2.97*	1.40	0.55	0.22	0.79	0.95	0.14	0.24	0.26	-1.91*	-0.47	-0.12	-1.77*	-0.74	0.55	1.36	0.36
	(1.10)	(1.62)	(1.19)	(0.90)	(1.31)	(0.98)	(0.80)	(1.20)	(0.83)	(0.77)	(1.09)	(0.85)	(0.58)	(0.98)	(0.59)	(1.17)	(1.56)	(1.29)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
b. *Comunero*																		
Killed (0/1)	-0.47	-0.41	-0.75	0.10	-0.54	-0.19	-0.87*	-0.56	-0.55	0.17	-0.36	0.29	-0.06	0.19	-0.15	-0.22	0.34	-0.35
	(0.62)	(0.86)	(0.66)	(0.47)	(0.62)	(0.51)	(0.47)	(0.67)	(0.46)	(0.41)	(0.64)	(0.46)	(0.34)	(0.52)	(0.33)	(0.65)	(0.83)	(0.73)
Killed (0/1) x Comunero	0.80	0.67	0.97	0.07	0.05	0.14	1.88	-1.68	1.38	0.27	-2.77*	-1.13	-0.42	-0.88	-0.31	3.36**	3.09	3.76**
	(1.56)	(2.18)	(1.70)	(1.50)	(2.39)	(1.62)	(1.51)	(2.01)	(1.53)	(1.19)	(1.56)	(1.23)	(0.91)	(1.20)	(0.96)	(1.58)	(2.51)	(1.72)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650
c. Native																		
Killed (0/1)	-1.22*	-0.44	-1.21*	-0.24	-0.77	-0.36	-1.11**	-0.37	-0.61	0.37	-0.16	0.70	0.22	0.72	0.33	0.03	0.59	0.22
	(0.65)	(0.88)	(0.70)	(0.49)	(0.63)	(0.54)	(0.54)	(0.77)	(0.52)	(0.47)	(0.73)	(0.53)	(0.44)	(0.72)	(0.44)	(0.71)	(0.85)	(0.81)
Killed (0/1) x Native	2.42**	0.26	1.51	0.99	0.75	0.48	1.05	-0.95	0.42	-0.58	-1.26	-1.47**	-0.93	-1.96**	-1.46**	-0.09	-0.09	-0.80
	(1.06)	(1.49)	(1.15)	(0.86)	(1.22)	(0.87)	(0.79)	(1.06)	(0.85)	(0.66)	(0.97)	(0.71)	(0.58)	(0.84)	(0.59)	(1.17)	(1.70)	(1.21)
R^2	0.09	0.09	0.09	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.10	0.10
Observations	115492	115492	115492	120595	120595	120595	114418	114418	114418	111480	111480	111480	116085	116085	116085	124650	124650	124650

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

Table 1.17: Identity Interacted Dummy

	Local/Neighbour			Ethnic/Race			Religion		
	All	A	T	All	A	T	All	A	T
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
a. Share									
Killed (0/1)	-1.20*	0.38	-1.19	0.44	-0.17	0.22	0.95	-0.53	1.13*
	(0.70)	(0.87)	(0.75)	(0.41)	(0.66)	(0.49)	(0.64)	(0.80)	(0.63)
Killed (0/1) x Share	-1.90	-3.68**	-1.39	0.71	2.27	0.28	0.88	1.66	0.82
	(1.16)	(1.75)	(1.25)	(0.92)	(1.44)	(1.05)	(0.85)	(1.19)	(0.86)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
b. *Comunero*									
Killed (0/1)	-1.62**	-0.48	-1.47**	0.26	-0.14	-0.09	1.33**	0.28	1.49***
	(0.63)	(0.83)	(0.68)	(0.46)	(0.60)	(0.51)	(0.52)	(0.70)	(0.52)
Killed (0/1) x Comunero	-5.54***	-8.01***	-4.60**	7.07***	11.89***	6.07**	-0.07	-2.17	0.02
	(2.07)	(3.00)	(2.28)	(2.30)	(3.67)	(2.51)	(1.37)	(2.05)	(1.47)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279
c. Native									
Killed (0/1)	-1.37*	-0.06	-1.41*	0.67	-0.04	0.36	0.62	-0.46	0.90
	(0.70)	(0.93)	(0.79)	(0.53)	(0.71)	(0.63)	(0.60)	(0.74)	(0.58)
Killed (0/1) x Native	-1.75	-2.97*	-1.00	0.21	2.27	-0.05	1.94**	1.71	1.55**
	(1.20)	(1.77)	(1.29)	(1.16)	(1.70)	(1.22)	(0.80)	(1.12)	(0.78)
R^2	0.13	0.13	0.13	0.20	0.20	0.20	0.09	0.09	0.09
Observations	126279	126279	126279	126279	126279	126279	126279	126279	126279

Notes: [1] Numbered columns refer to the violence variable used in the estimation by each social trait. *All* holds for overall violence, *A* for army-led killings, while *T* for terrorist-led killings. [2] Panel a. uses the number of impressionable years affected by the violence, while panel b. re-categorizes it to a dummy variable denoting at least one impressionable year affected by conflict. Panel c uses the number of people killed during the period of the impressionable years scaled by the district population in 1993. [3] All regresions include dummies for year, age, district of birth, district of residence, and a set of province-especifict trends. Additional controls: years of schooling, a dummy for males, a dummy for native speakers and a dummy for residents in the coast region. [4] Standard errors clustered at the district of origin level. [5] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

As for trust, native people show a higher effect in response to violence during their impressionable years when it comes to trust the political parties (as shown in table 1.16, panel (c)).

In sum, the results of this sub-section point towards an heterogeneous effect of violence during the impressionable years depending on the degree of connection to the indigenous dimension. This effect is important for the group of variables about identification. The general result that emerges is that violence, and in particular, army violence, harmed people's identification with locals, and on the other hand strengthened the feeling of identification with ethnic and race groups.

1.6 Potential Channels and Further Avenues for Research

Hitherto this chapter has found an statistically significant effect of violence during the impressionable years over the levels of trust on some institutions and the degree of identification with groups of population. The effect varies depending on the indigenous origin of the individuals. Although an historical account of the violence period could shed light on the reasons for this effect, it is necessary to build a theoretical explanation for the results. In this section I provide a discussion in this regard without aiming to build a fully formed theoretical framework. In fact, I draw from Bauer et al. (2016) to guide the discussion, and explain the potential channels they consider in their meta-analysis.

Bauer et al. (2016) propose three theoretical approaches from which future research could start explaining the emerging body of statistical results: (i) changes in economic constraints, (ii) changes in parochial norms and preferences, and (iii) changes in general preferences and other psychological explanations.

1.6.1 Economic constraints

The first approach seems the most immediate to adopt in the economics discipline. It could be possible that exposure to violence at some point in life alters budget constraints and makes optimal to trust or trust less certain institutions, even more in the case if identity. In perhaps a relevant attempt to build a theoretical approach to explain violence participation, Guardado

(2015) proposes a model in which the decision to participate in violent activities varies depending on the communal relevance of the district where Peruvian individuals reside. In her model, individuals who live in districts with a high degree of communal land concentration (measured with tenancy) have an smaller probability to engage in violent activities during coffee prices shock. The interpretation that the author provides rely on a model of labor decisions that considers different contract arrangements. Living in a district with a high concentration of communal land offers the possibility to risk-share in the presence of the price shock, therefore the need to engage in alternative activities, such violence, decreases. Applied to the context and target of this chapter, a model like this would account for the possibility that changes in the degree of identification, for instance, alter the optimal contract that allows individuals to cope with the shock. If the individuals feel less identifies with their locals, would that make it more costly to engage in a risk-sharing contract?

1.6.2 Parochial norms

The CVR interpreted the ethnic and race factor as a dimension that crossed both combat fronts in the Peruvian case. According to the CVR, the conflict in many cases provided the opportunity to exacerbate latent communal conflicts (over demarcation, public good provision, etc.). In this sense, a theoretical explanation that relies on the alteration of parochial norms would try to understand how the exposure to violence deepened the frictions among communities whereas at the same time increased the sense of identification with individual's own community. As mentioned in the section about communal land (1.2.2) these types of communities still engage on certain types of local cooperation within the community. However, it is also know that they cooperate very little between them. The agricultural census of 2012 indicates that among these communities, a mere 3% makes businesses with other communities. Therefore, in this context, a theoretical framework should would attempt to understand how violence changed the degree of integration within the local communities.

1.6.3 General preferences and other psychological explanations

As noted by the CVR, the Peruvian conflict had devastating effects on individuals directly affected by the violence. There is a series of case studies which highlight a general sense

of sadness and disregard for any type of social interaction among the population with the highest exposition to the violence. The case studies presented in the CVR report also manifest that these effects were persistent. In that regard, a potential theoretical framework in this context would look to identify whether further generations of people with heavy exposure to violence also evidence signs of disregard and general discomfort.

1.7 Conclusions

Wars can have persistent effects that go beyond the evident physical destruction. Among the non-physical effects, beliefs constitute an important topic to analyze. Its relevance relies on their connection to culture and institutions, which ultimately explains the degree of development experienced by some nations.

In this chapter I evaluate the effect of a 20 years armed conflict on individual's beliefs about trust and identity. The way it affected is through its shocking influence during a critical moment in people's life, the impressionable years.

The baseline results indicate that violence experienced during the impressionable years at the district of birth of the individual had a negative effect on the degree of trust over some institutions: the police, the judiciary system and the political parties. The estimated effect, nonetheless, is small. A one standard deviation increase in the number of impressionable years exposed to overall violence reduces the probability that an individual trusts the police by 0.24 percentage points, the judiciary institution by 0.27 percentage points and the parliament by 0.19 percentage points.

The baseline results for the variables on identity show that the exposure of at least one impressionable year to overall violence reduces the probability that an individual identifies himself or herself with his or her local co-inhabitants by 2 percentage points, while it increases the probability that he or she finds identification with a religious group by 1.32 percentage points.

Using a different period of exposure, early childhood, offers different results. The effect on trust is positive, however also small, whereas there is no detectable effect on identity. This in itself is an interesting result that indicates that exposure at different stages in life to traumatic events may lead to opposing consequences in the formation of people's beliefs.

In the context of the Peruvian conflict, the ethnic and race component played an important role in spite of not being the fundamental cause for the conflict. This is particularly relevant for the variables addressing identity. Individuals who own an agricultural plot which is embedded in a communal tenancy exhibit higher negative effects on the probability of identification with locals, while higher levels of identification with their ethnic group or race.

The last result ultimately confirms the speculations of the Commission for the Truth and Reconciliation for the Peruvian context that the ethnic and racial component, although explicitly neglected during the conflict period had a considerable importance in the process of understanding the consequences of the war.

In order to further understand these consequences further research needs to build a theoretical framework that mixes changes in the economic constraints, changes in the parochial preferences and in the general preferences after the individual was exposed to the violence.

2 The Determination of Democratic Beliefs Under Uncertainty

2.1 Introduction

"Democracy is the worst form of government, except for all those other forms that have been tried from time to time" is an expression attributed to Winston Churchill, which depicts a subtle and persistent hesitation about the benefits of democracy that still permeates the political debate in some regions of the planet.

Despite the claims about its positive effect on income (Acemoglu et al.; 2014)[22], why democratic regimes are still not the norm? Why only 35 out of the 164 countries covered by the Polity-IV index in 2013 are categorized as full democracies? Or put differently, why 45 of these 164 countries have a political regime that ranges from anocracy to autocracies or failed states?

Taking an specific context, why only 19.6% of Peruvian individuals surveyed between 2007 and 2012, think that democracy is important? In 2007 Peruvians had been living in democracy for 6 years since the last autocratic regime that left power amid blatant allegations of corruption and a widespread economic crisis. After 2001 the country grew vigorously, which in a way seems to validate the point that democracy is at least associated with economic growth, yet, why only 19.6% of Peruvian inhabitants deem it important?

The ability to judge something as *important* emerges from individuals and their beliefs. Naturally, the sum of individual judgments, or beliefs, define policies and any nation's fate, but still, it is the individual and his or her beliefs what determines the fate of societies.

There is a growing literature that analyses the formation of beliefs. Among this literature, in a recent and influential study, Giuliano and Spilimbergo (2014) find that US individuals who grew during a recession tend to believe that success in life depends more on luck rather than effort; support more government redistribution; and are less confident in public insti-

[22]The determination of democracy, on the other hand, is a long-standing question in economics literature. Acemoglu and Robinson (2006) revived the discussion on the determinants of democracy. Acemoglu et al. (2008), suggest that controlling for variables that mutually affect democracy and income, the apparent relationship found in many studies (Dhal; 1971, Huntington; 1991) disappears. Bruckner and Ciccone (2011) presented evidence that rainfall shocks led to a window of opportunity for a democratic transition. The consolidation of democratic regimes has received some attention too. Alemán and Yang (2011), for instance, highlights the importance of institutional and political factors (oppositionist social mobilizations) beyond the only economic factors like income inequality. Svolik (2008) distinguishes between democracies that survived because they consolidated and democracies that survived by some favorable circumstances.

tutions. The effect the authors find is long-lasting.

But perhaps the key feature in Giuliano and Spilimbergo (2014)'s analysis is that recession during a period in the individual's life, the impressionable years, carries the effect, and not experiencing recession at any other point in time. The impressionable years period is the period when the individual is aged between 16 and 25 years, and according to psychologists, when a he or she develops the set of attitudes and beliefs which will shape his or her life in the future. This is the main reference, and guideline I will use in the analysis proposed in this chapter[23].

On the other hand, climate variability is generally linked to economic risk in rural societies. Agricultural productivity is crucially related to climatic volatility (Mendelsohn; 2007). However the range of outcomes linked to weather volatility also cover health (Gallup and Sachs; 2001), time preferences (Galor and Ozak; 2016), conflict (Miguel and Satyanath; 2011; Barron et al.; 2014; Hsiang et al.; 2013) and more generally, economic growth (Dell et al.; 2012).

While the literature that links rainfall shocks to beliefs is relatively new. In particular, some researchers have placed special interest on the effect of risk over beliefs. Durante (2010) models the effect of rainfall risk on trust at community level for several countries in Europe. The author finds that regions characterized by higher year-to-year rainfall and temperature variability display higher levels of trust, which is consistent with a theory of insurance through geographical differentiation. This result is statistically significant for the two measures of risk evaluated separately: historical and current, but remains only for historical when both types of risk are included. Historical variability is the monthly rainfall for the period 1500-1750, and current variability covers the period 1900-2000[24].

Durante (2010)'s explanation for the result is that communities where historically rainfall has been more volatile developed a bigger sense of cooperation. Economically this emerges

[23]In this study, the authors do not focus their attention on democratic beliefs, however Giuliano and Nunn (2013) provide evidence on the transmission of democratic values through generations. Using a classification of democratic tradition based on the language and dialects of current population, the authors show that past experience with local-level democracy is associated with more supportive beliefs of democracy today.

[24]Temperature volatility has also been associated to political unrest. Durante et al. (2015) find that heat stress in spring, as a proxy for negative income shocks, increase the likelihood of violence against Jews, but only during pogrom waves that are triggered by political turmoil. Hsiang et al. (2013) find that weather volatiliy is related to human conflict.

as a coordinated response to a problem of insurance. Could democratic preferences also be affected by risk? It is possible that democracy, as a collective action, represents a sort of mechanism to insure average people's interest. It is also possible that uncertain environments, where anybody could get a bad shock, develops a deeper sense for redistribution, which in a way has a democratic flavor. I frame the shaping of democratic beliefs within the idea that risk entails cooperative behavior, and democracy acts as a cooperation instrument.

The macro literature the links rainfall shocks to democracy includes the work of Bruckner and Ciccone (2011) which shows that negative rainfall shocks are followed by an improvement of the democratic conditions among some African countries. This effect operates through social uprising due to the deterioration of the local economic conditions which lead to the desire for a political change in favor of democracy. The reason why democracy is preferred, I would argue, links to the theory proposed by Durante (2010): tightening economic conditions due to an income shock or income uncertainty motivates the emergence of cooperative systems, where individuals have more incentives to trust each other and create a form of social capital. In the same line, Haber (2012)suggests that countries located in latitudes with high climate volatility tend to prefer autocratic regimes to democratic ones.

In this chapter I intend to provide micro evidence to test whether the hypotheses of Bruckner and Ciccone (2011) and Haber (2012) hold within the same country. In doing so, I draw inspiration from the studies of Giuliano and Spilimbergo (2014) and Durante (2010), and propose an empirical model to assess the role of uncertainty during the impressionable years over the democratic beliefs. For the measurement of risk I follow the well known study of Paxson (1992), where the author measures the extent to which farmers in rural Thailand are able to use savings to smooth consumption in response to unexpected shocks to income.

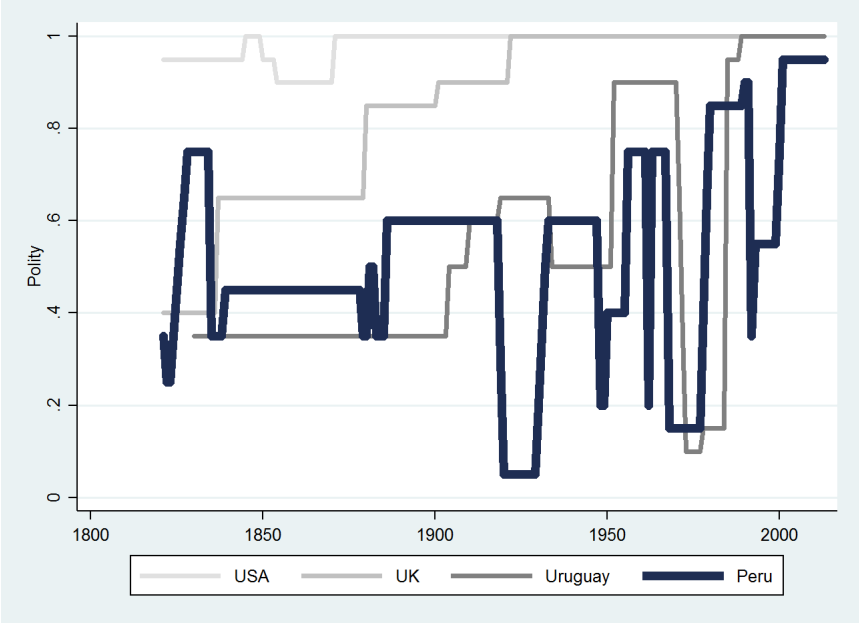
In this chapter, therefore, the focus of the analysis is on preferences for democracy, or beliefs about democracy and its relationship with uncertainty. The idea behind is that the prospect of uncertainty (income, consumption, or wealth) motivates individuals to become more cooperative and therefore more confident about a system that privileges a shared common interest. More importantly, is it uncertainty during the time the individual is being surveyed what matters? Or is the uncertainty she experienced during her formative years?

But uncertainty may not be the only explanation. The combination of a risky environment

and anti-democratic regime that somehow manages to stabilize the economy could spur a sense of frail democracy, incapable of providing economic insurance. In this chapter, I intend to explore whether the exposure to authoritarian regimes combined with some degree of economic uncertainty helps to explain democratic beliefs.

The analysis is focused in Peru, whose democracy can be described as troubled. Although established as a republic since 1821, Peruvian political history has been characterized by an alternation between democratic and authoritarian regimes. This, however, is a major feature of democracies in Latin America, where the consolidation of democratic regimes seems to be always on the verge of turning backwards. Figure 2.1 shows the normalized Polity IV index for the United Kingdom, the United States of America, Uruguay and Peru. The average normalized Polity IV index for Peru for the period 1821-2013 is 0.53. Uruguay, as another example of Latin American politics, has a slightly better performance with an average index of 0.55. It is clear, nonetheless that both countries totally under-perform USA (0.98) and UK (0.85).

Figure 2.1: Polity Index for the period 1821-2013



Notes: [1] Data comes from the Polity IV Individual Country Regime Trends, 1946-2013 report, from http://www.systemicpeace.org/polity/polity4.htm. [2] The chart shows the normalised trend of variable polity2.

In that sense, the Peruvian context offers a useful framework to test for preferences in

democracy. Anecdotal evidence suggest that in Peru many people are still fond of past authoritarian regimes and the reasons vary from government to government[25]. Would it be possible that the experience of uncertainty under authoritarian times passed unnoticed precisely because of the troubled context?

To explore this question I focus on a rural sample of individuals of different ages and locations from 2002-2012. Rural population in Peru is still around 27% and agriculture plays an important role in Peruvian economy, in 2012 agricultural valued added represented 7.2% of total valued added in the country. By focusing on rural individuals, I expect rainfall to have a larger importance than for urban individuals. Rural society in Peru is heavily dependent on agriculture, where 43% of working age individuals are engaged in any agricultural activity.

In measuring democratic preferences I develop 9 measures from the surveys grouped in four categories: (i) the belief that democracy is important, (ii) beliefs on the performance of the current democratic regime, (iii) the belief that despite its imperfections, democracy is the best option, and (iv) it is the preferred type of government. To measure economic uncertainty I rely on rainfall data constructed at district level.

To preview the results, I find no statistically significant effect of uncertainty during the formative years on democratic beliefs. Rural individuals exposed to either a rainfall shock or rainfall uncertainty during their impressionable years did not change their democratic beliefs. The (lack of) result is consistent through the 9 measures of democratic beliefs. The exposure to uncertainty which coincided with an authoritarian regime does not affect the democratic beliefs either.

This result is robust to a different definition of rainfall shocks, a narrower definition of rural individuals, and after controlling for the degree of authoritarianism that the individual experienced. Allowing for different formative years does not change the main result. When I evaluate current uncertainty, as additional check, this type of uncertainty seems to explain

[25]In recent presidential elections, held on June 5th 2016, 49.88% of the electorate population voted for Keiko Fujimori, the daughter of the last dictator, Alberto Fujimori. The electoral campaign was charged with massive manifestations against a likely return to a government of the characteristics of Alberto Fujimori. As in any campaign, accusations went from one side to the other, however Keiko Fujimori's contenders all pointed to her lack of democratic credentials. Yet, almost half of the electorate population preferred her to Pedro Pablo Kuczynski, who finally won the presidency with 50.12% of the votes.

part of the belief on the importance of democracy, however this result may be a consequence of not controlling for current political trends.

The chapter is organized as follows: section 2.2 details the data and construction of the variables to measure democratic beliefs and economic uncertainty. Section 2.3 presents the main evidence and some variations of the main equations. Section 2.4 provides a major robustness check: allowing for different formative years. Finally, section 2.5 concludes.

2.2 Data

In the estimations I use household and rainfall data from Peru. The source for the household data is the National Household Survey (ENAHO in Spanish) that collects information from a nationally representative sample of households on a yearly basis since 1998. The organization in charge of collection is the National Institute of Statistics (INEI in Spanish). However, I restrict the analysis to individuals living in a rural environment during the period 2002-2012, the period for which the democracy questions are available and, depending on the measure, I am left with observations between 15000 and 52000. The unit of analysis is any individual aged 18+ that answered the question on democracy living in a rural environment. Section 2.2.1 provides detailed information on the construction of the dependent variables.

I match individual's beliefs in the period 2002-2012 to rainfall data using the district identifier. For the construction of the measures on uncertainty based on rainfall shocks, I use data from the Climatic Research Unit (CRU) at the University of East Anglia. This generated data set ranges from January 1901 to December 2012. The level of detail is 0.5x0.5 degrees. For the measurement of uncertainty I can compute measures for all districts. This is explained din detail in section 2.2.2.

2.2.1 Household Data

To investigate the determination of democratic preferences I group democratic beliefs in four types of questions: (i) how important is democracy, (ii) as preferred type of government, (iii) whether it works, and (iv) even with its imperfections, it is the best alternative. All the questions are intended to qualify the degree of satisfaction of the individual with respect to democracy. Individual satisfaction with democracy could mask what the person think of

his current government. This is a possibility for someone who has lived all her conscious life under a democratic regime, for whom the question about satisfaction with democracy can be understood as her satisfaction with the government. I will argue that even if that is the case, expressing satisfaction with the government means by extension expressing satisfaction with the political system that hosts the government.

Moreover, and as figure 2.1 shows, the experience of non-democratic or authoritarian regimes is still very recent in Peru. For instance, someone 50 years old interviewed in 2007 would have lived under both non-democratic and democratic regimes during her impressionable years. For this hypothetical person her impressionable years took place during the period from 1975 through 1982, where the sub-period 1975-1979 was marked by the rule of dictator general Bermudez, while the sub-period 1980-1982 president Belaunde was in charge after dictator Bermudez resigned in 1979 and called elections. Clearly this person experienced two very different political systems and even if she expresses dissatisfaction with the government in 2007 that would be an expression of dissatisfaction with a democratic type of government. Empirically, the cohort fixed effect controls for the experience of democratic and non-democratic regimes during the impressionable years while the survey year fixed effects control for any attitude towards the current government: if Humala's government is more popular than Garcia's one, this should be captured by the survey year fixed effect, leaving any remaining information to be associated with the perception of democracy I am precisely trying to capture[26].

The political history of Peru offers a lot of variation in this regard. Consequently, I assume that when individuals respond to the question about democracy they have an idea of what a non-democratic regime is. A question like this in a different context, the USA or the UK for instance, would be of little relevance since the individual there would have not had the chance of living under a non-democratic regime, which would cast even more doubts about what this person is actually responding. Peruvians Have been exposed to huge political variation, which provides the source of variation I exploit in the identification strategy.

[26]Empirically, the inclusion of year and cohort fixed effects –through the inclusion of age dummies at the same time of year dummies, rules out preferences associated with perception of the government, since some of the questions on democracy cover different government periods. Consider for instance, the question on the importance of democracy for the period 2007-2012. This question overlaps with two governments: Alan Garcia (2006-2011) and Ollanta Humala (2011-2016).

Leaving aside the consideration about the exposure to different political regimes, the wording of the questions is also worth discussing. The first group of questions (C.1) are concerned about the importance of democracy. Depending on the success of a democratic regime, individuals have enough information to ascertain whether living under democracy is important. If the political system changed but the individual sees no change in her situation, it is then valid to think that democracy represents nothing different for her.

The second group of questions (table C.2) directly inquires individuals about political preference between democratic and authoritarian regimes, and in this case the ambiguity between the perception about democracy and the current government is less important. This relates to the third group (table C.3) where once the surveyor has informed the individual that even if democracy is not perfect, inquires about its relevance as an option.

The last group of questions are related to understand whether democracy works (C.4), which fortunately kept the same wording from 2004 to 2012. The period 2002-2012 overlaps with three different governments: Alejandro Toledo (2001-2006), Alan Garcia (2006-2011) and Ollanta Humala (2011-2016). The inclusion of the year dummies controls for different responses about democracy that are associated with the different government regime.

In general, these government periods overlap through all the questions and I rely on the year dummies to control for any association between the expression of (dis)satisfaction about democracy and the (dis)satisfaction with the current government.

Appendix C shows the literal translation from Spanish of the survey questions, the answer alternatives coded as 1 and 0 and the corresponding averages. Although each group of question refers to a main topic related to democracy, the framing of the question sometimes changed within the group along the 2002-2012 period. Tables C.1 to C.4 show the categorization for each type of variable and its national and rural average. I dropped individuals reporting "Do not know", which represented around 6% of the total sample.

Although all questions retained their core object of study (importance of democracy, performance, etc) the framing changed through the years. In what follows I comment on the main features of every topic by reference period.

Table C.1 groups the three variables related to the belief that democracy is important. Through the period 2002-2006 the question was the same, however, the replacement of the

option *more or less* in the 2002-2004 sub-period by *enough* in the sub-period 2004-2006[27] altered the proportion of individuals responding that democracy has *a lot* of importance. The national average slightly decreases from 0.456 to 0.426, whereas the rural average changes from 0.348 to 0.328[28]. It is important to notice that within each period averages are very similar.

In the third sub-period, 2007-2012 the question and alternatives changed. In order to keep it comparable with the other two groups, this variable is re-categorized to take the value of 1 for individuals stating that democracy is *very important*. Including the option *Important* in the account increase the proportion of positive respondents to 0.826, instead of 0.196. The proportion of rural individuals who declared that democracy is *Very Important* from 2007-2012 is 0.131. Clearly this makes sub-periods relatively comparable, but I decided to study them separately. As suggested by Bertrand and Mullainathan (2001), small changes in the wording in the question could produce changes in the way people answer questions.

Beyond the differences in framing and averages between groups, the most noticeable fact from table C.1 is the low proportion of inhabitants who consider that democracy is very important. Less than half of the national sample has a very positive appreciation about the importance of democracy. Such disregard for the importance of democracy is even deeper for individual living in rural areas. Less than 40% of the rural inhabitants in the sample consider that democracy is very important.

The second group of questions relates to the preferred type of government. The individual is presented with a set of different type of governments and she has to choose one among them as the preferred option. The options are between democratic type of governments, authoritarian or military governments or a government ruled by "experts". The change in framing and alternatives for the questions is more evident here, as presented in table C.2. For the 2002-2006 sub-period, the democratic option is presented in third place, after the options of authoritarian or military regimes, before the final category: independent people[29]

[27]There is some overlapping between this two periods because the 2003 sampling covered some months of 2004. These months were January to April and were sampled under the 2003 survey design.

[28]If the *Do not know* category is re-coded as 0, the differences are more evident. More formally, a test for the differences in national means yields a t-stat of -8.80, which rejects the null hypothesis of equal means. The same test for the rural averages also rejects the null hypothesis of equal means with a t-stat of -4.89

[29]Understood as experts, professionals as opposition to politicians.

regimes. The question for the 2007-2012 period, is set in a way that obliges the individual to prefer a democratic regime to an authoritarian or to be indifferent between regimes. For both periods, the dependent variable takes the value of one when the individual states that democracy is more convenient, or if they agree that democracy is always preferable. Clearly the alternative that refers to a democratic regime is stronger in the second sub-period due to the mention *is always preferable*. And that may explain the difference in the average for both sub-periods.

60.5%[30] of the individuals in the first period prefer a democratic type of government when presented with other options. Again, the rural average is smaller, 55%. In the second period, 56.5%[31] think that a democratic government is always preferable. Such perception is smaller for rural individuals, who only 49.1% would prefer a democratic regime.

The third type of questions , shown in table C.3, looks like one that Winston Churchill would have answered positively. It evaluates the preference for democracy even if it is not the best political system. In this case, the individual is asked for a degree of agreement with the statement that democracy may be the best option in spite of its flaws. The difference in the framing between periods is also evident for both, the question and the alternatives. In the first period, 2002-2004 the question stated that although democracy can cause problem it is the best option. The alternatives *Much* and *More or less* were considered as positive agreement. The second period, 2004-2006, offered a question that asked for agreement on the fact that democracy has imperfections but it is the best option available. Again, the first two choices, *Much* and *Enough* were considered as agreement with the statement of the question. The household surveys between 2007 and 2012 did not have this question.

The average difference between the two periods is large. In 2002-2004 65.5% of the national individuals would agree that democracy is the best option, whereas in the second period, less than half of the individuals, 45.9% agreed that democracy is the best option. This difference arises as the category *little* grew from 17.26% in the first period to 29.14% in the second period. The rural averages, again, are smaller. The percentage of individuals agreeing with the statement in the first period was 53.4%, while in the second period it was

[30] 7.4% preferred an authoritarian government, while 14.4% preferred a military government.
[31] 15.1% preferred an authoritarian government.

34.3%.

The final group of questions, grouped in table C.4 explore the belief that democracy is working. Although the question was the same through the whole period, the choices were different for two periods: 2002-2004 and 2004-2012. In the first period period the first choice, *A lot*, may accumulate most of the positive responses, whereas in the second period, *Very well* and *Well* have a clear positive meaning. I re-coded it to make the first two categories a positive statement on the belief that democracy works.

The averages, however, differ quite importantly from one period to the other. The national average was 55.3% in 2002-2004 but it was just 35.9% in 2004-2012. Interestingly the rural average in the first period, 54.9%, is not too different from the national average. Actually, in the second period the rural average is higher than the national mean: 40.3%.

In sum, I have 9 variables to test the effect of uncertainty on democratic beliefs. This 9 variables are grouped in four main categories regarding (i) the importance of democracy, (ii) as preferred type of government, (iii) as best option in spite of its imperfections, and (iv) its performance. The general picture that emerges from these tables is one of generalized distrust for democracy, especially in the rural areas, which constitute the sample for the analysis, as explained above.

To sum up, I decided to treat every question framework separately. As showed in the descriptive tables, small changes in the wording of the questions ans alternatives alter the way people respond to the questions. This is a concern widely recognized in the literature and noted particularly in Bertrand and Mullainathan (2001). All the questions, however, have been re-coded to reflect a positive attitude towards democracy, therefore I will interpret any effect of rainfall shocks and uncertainty in this context.

It is probably the systematic alternation of democratic and authoritarian regimes what explains the low level of acceptation of democracy as a relevant system. Although Peru gained independence from Spain in 1821, it was not until 1872 when Manuel Pardo took oath as president of Peru in the first elections where a civil servant not related to the army won the presidency. This cycle was interrupted by the war against Chile in 1879, when a succession of self-proclamations and parliamentary designations decided who would be the head of the state. The next period of elections was in 1886, in the aftermath of the war. Since

then, the alternation of democratically elected presidents and generals who seized the power through a coup d'etat was the norm. 1980 was the last year when a general non-elected was president. 1992, however, marked the beginning of a new type of authoritarianism, when the elected president Alberto Fujimori shut the parliament. His presidency lasted until 2000 when he fled to Japan after scandalous allegations of corruption in his government. The last, and perhaps the longest, democratic period started in 2001 and continues to this day.

2.2.2 Rainfall Data and the Measurement of Uncertainty

I construct two measures of uncertainty using rainfall data: the standard deviation of rainfall, and the anomaly. To calculate them, I use monthly rainfall data from the Climatic Research Unit (CRU) from the University of East Anglia[32]. The data is available from January 1901 to December 2012 at a resolution of 0.5x0.5 degrees grid. All rainfall variables are constructed at grid level and then are averaged at district level with a weight depending on how much of the district extension is on the grid[33].

In concrete, I construct four main rainfall variables associated to uncertainty and risk: (i) a measure of uncertainty during the impressionable years of the individual, σ_{iy}, (ii) rainfall anomaly (rainfall deviation with respect to an historical average) during the impressionable years, D_{iy}, (iii) current (during the year of the survey) uncertainty, σ_{10}, and (iv) current rainfall anomaly, D_{10}. I begin with the discussion of the last two measures.

The reason to include rainfall uncertainty and rainfall anomaly rests upon the logic followed by Paxson (1992). Although the main measure the author chooses for her analysis is the measure of rainfall anomaly, I focus the analysis on the measure of uncertainty, σ.

Current uncertainty, σ_{10}, is a year-to-year measure over a period of 10 years. The calculation process is as follows:

1. I estimate a 10-years monthly average for rainfall in grid j and month m: $\bar{R}_{jm} = \sum_{k=y-1}^{y-10} \frac{R_{jkm}}{10}$

[32] http://www.cru.uea.ac.uk/

[33] I use information of the district borders (polygons contained in a shape file) and overlaps the district borders to the grid of the world map at a resolution of 0.5x0.5 degrees. I then calculate the weight of each grid for the calculation of the district average. The weight is calculated based on the number of points the district polygon crosses the 0.5x0.5 grid.

2. The anomaly is the current deviation of the rainfall in grid j in month m and year y with respect to the mean estimated in 1.: $D_{jym} = [R_{jym} - \bar{R_{jm}}]$

3. I average the last 12 months anomalies in order to get an annual estimate of current anomalies for grid j and year y, which I call D_{jy10}

4. I map the estimate of the anomaly in the grid j to the district i[34], and calculate the measure at the district level, which for ease of notation I call D_{10}.

5. The second and more important measure, uncertainty, is the standard deviation of rainfall over the 10 years period. I first estimate it as a monthly measure for grid j and month m in year y which takes as reference the 9 years preceding y and year y itself: $\sigma_{jym} = \sum_{k=y}^{y-9} \frac{(R_{jym} - \widehat{R}_{jm})^2}{9}$

6. Then I average it across the last 12 months, map the grid to the district, and define the district measure (ignoring district and year indexes): σ_{10}

It is important to notice that the means used in the estimation of the anomaly and the uncertainty measure are different. The anomaly measure, D_{10}, is estimated using an historical average, $\bar{R}$, which does not consider year used for the estimation of the anomaly; whereas the uncertainty measure, σ_{10} is simply the standard deviation with respect a mean within the 10 years period, $\widehat{R}$.

The measures of uncertainty and anomaly during the impressionable years are constructed using the measures of current uncertainty and anomalies previously calculated. In particular anomaly and uncertainty during the impressionable years are constructed over the annual averages of each of the impressionable years.

1. Anomaly during impressionable years is defined as: $D_{iy} = \sum_{k=18}^{25} \frac{D_{10,k}}{8}$

2. Uncertainty is defined as: $\sigma_{iy} = \sum_{k=18}^{25} \frac{\sigma_{10,k}}{8}$

Table 2.1 presents the basic statistics for the anomaly and uncertainty measures estimated for the current period of the survey and the impressionable years broken by national and rural individuals.

[34] As described in footnote 33

Table 2.1: Uncertainty and Anomaly Basic Stats

Variable	Obs	Mean	S.D.	25th Perc.	75th Perc.	Number of Districts
National						
D_{iy}	609319	0.001	0.111	-0.044	0.049	1296
D_{10}	635114	-0.043	0.258	-0.188	0.096	1266
σ_{iy}	609319	**0.429**	0.216	0.280	0.557	1296
σ_{10}	635114	**0.514**	0.227	0.337	0.660	1266
Rural						
D_{iy}	230191	0.000	0.103	-0.041	0.046	1118
D_{10}	242027	-0.034	0.243	-0.171	0.100	1094
σ_{iy}	230191	**0.455**	0.229	0.289	0.584	1118
σ_{10}	242027	**0.537**	0.240	0.363	0.681	1094

As I stated previously, the focus of the analysis is on the measures of uncertainty, σ_{iy} and σ_{10}. Uncertainty during the impressionable years is smaller than current uncertainty because it is an average over a period of 8 years, the impressionable years. For the national sample current uncertainty is 0.514 and during the impressionable years is 0.429 on average. In the rural sample both types of uncertainty are higher. Current uncertainty is 0.537 and uncertainty during the impressionable years is 0.455. For an individual living in rural areas the distance between the 25th and the 75th percentiles is 0.295.

Uncertainty and anomaly during the impressionable years are matched to the district of birth of the individual. Unfortunately the survey does not allow to identify the residency of the individual during different periods. Therefore I assume that experiencing uncertainty or rainfall anomaly in the birth district accounts for the same experience during the impressionable years[35]. Current uncertainty and anomaly, on the other hand, are matched to the current district of residency of the individual.

The correlation between the measure of rainfall anomaly (D_{iy}) and uncertainty (σ_{iy}) is -0.029 and is statistically significant at 1%. I do not expect this, however to be a main source of concern since the rainfall anomaly ranges large negative values to large positive values while the second measure can only take positive values by construction. The only case where this could be a source of concern is if I used only positive values of D_{iy} or only negative values of D_{iy} in the same specification where I use σ_{iy}[36]. This is not the case, since section 2.3.3, where

[35]Mobility among rural individuals is relatively low, 64% of the individuals report living in the same district they were born. At a larger administrative unit, province, 73% of the individuals are living in the same province that they were born.

[36]Even in this case, the correlation between these two measures rises to 0.40 when the rainfall anomaly is constrained to only positive or negative values

I distinguish between positive and negative shocks, I still include both types of shocks but re-define D_{iy} to take the value of 1 for negative shocks and 0 for positive. The correlation between this dummy measure and σ_{iy} is close to zero.

2.3 Evidence

This section presents the empirical estimations to track the effect of uncertainty and rainfall anomaly on the democratic beliefs of rural individuals. Most of the evidence is about uncertainty during the impressionable years, and different ways to approximate such relationship. The section closes with the estimates using current uncertainty and rainfall anomaly.

2.3.1 Main Evidence

Democratic beliefs are linked to uncertainty and rainfall anomaly through equation 2.1. I use a pooled estimation with several controls. The identification of uncertainty using rainfall data matched to the district of birth level provides a reliable source of exogenous variation for the estimation.

$$y_{itr} = \alpha_0 + \alpha_1 D_{iy}^{ro} + \beta_1 \sigma_{iy}^{ro} + \tau_t + \theta_a + \gamma_r + \gamma_{ro} + \gamma_r \tau_t + \Gamma X_{it} + v_{it} \tag{2.1}$$

Where y_{itr} is the democratic belief of individual i in time t who currently lives in district r. That belief is determined by the average anomaly, D_{iy}, and the measure of uncertainty σ_{iy} during their impressionable years matched at the district of birth ro. I also control for year fixed effects, τ_t, age fixed effects θ_a, current district fixed effects, γ_r, district of birth fixed effect, γ_{ro}, the interaction of γ_r and the year dummies τ_t. Finally, I also include a set of characteristics: age, male, household size, education level (no education as the base category), marital status, informal worker[37], unemployed worker and household income equivalent[38]. Table 2.2 presents these results.

The number of observations in each estimation is never smaller than 15,000. Each column of the table reports the estimated coefficients for each of the 9 measures of democratic

[37] Informal worker is the worker with no social contributions.

[38] Following Deaton (1997) I assume that children aged 0 to 4 are equal to 0.4 adults and children aged 5 to 14 are equal to 0.5 adults

Table 2.2: Democratic Beliefs and Uncertainty During Impressionable Years

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
D_{iy}	0.021	-0.045	0.013	0.000	-0.007	-0.004	0.013	-0.020	-0.012
	(0.047)	(0.041)	(0.021)	(0.050)	(0.046)	(0.050)	(0.023)	(0.026)	(0.027)
σ_{iy}	0.028	-0.026	-0.022	0.015	-0.050	0.029	0.046	-0.036	-0.055*
	(0.051)	(0.053)	(0.025)	(0.054)	(0.061)	(0.056)	(0.031)	(0.033)	(0.032)
Controls									
Male	0.018**	0.025***	0.035***	0.030***	0.028***	0.015	-0.006	0.031***	0.118***
	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.005)
Household Size	0.003	0.003	0.001	0.005**	0.003	0.009***	-0.000	0.005***	0.005***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Primary	0.057***	0.051***	0.051***	0.058***	0.025**	0.035***	-0.034***	0.080***	0.136***
	(0.010)	(0.010)	(0.004)	(0.010)	(0.011)	(0.011)	(0.005)	(0.006)	(0.006)
Secondary	0.119***	0.099***	0.120***	0.101***	0.089***	0.037**	-0.063***	0.145***	0.187***
	(0.014)	(0.014)	(0.007)	(0.014)	(0.015)	(0.015)	(0.007)	(0.009)	(0.008)
Technical	0.221***	0.195***	0.184***	0.166***	0.189***	0.060**	-0.077***	0.189***	0.249***
	(0.028)	(0.025)	(0.014)	(0.025)	(0.026)	(0.029)	(0.013)	(0.018)	(0.016)
University	0.308***	0.190***	0.212***	0.107***	0.292***	-0.003	-0.086***	0.208***	0.236***
	(0.036)	(0.037)	(0.020)	(0.039)	(0.036)	(0.042)	(0.018)	(0.025)	(0.022)
Married	0.007	0.005	0.002	0.019*	0.011	0.013	-0.008	0.024***	0.043***
	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.005)
Informal Job	0.015	0.002	-0.005	-0.008	0.002	0.002	-0.004	-0.012**	0.013**
	(0.010)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.005)
Unemployed	0.037	0.022	0.007	-0.034	-0.026	-0.005	-0.031*	0.024	0.029
	(0.028)	(0.028)	(0.014)	(0.029)	(0.031)	(0.028)	(0.016)	(0.018)	(0.018)
Household Income	0.019***	0.033***	0.018***	0.020***	0.028***	0.029***	-0.007**	0.023***	0.025***
	(0.007)	(0.007)	(0.003)	(0.008)	(0.007)	(0.008)	(0.003)	(0.004)	(0.003)
R^2	0.362	0.364	0.238	0.365	0.378	0.316	0.298	0.308	0.285
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\gamma_r x Year$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_{ro}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the district of birth; *significant at 10%, **significant at 5%, ***significant at 1%.

beliefs. Columns 1 to 3 belong into the group of democratic beliefs related to importance of democracy. Columns 4 and 5 are the beliefs that democracy is the best option in spite of its imperfections. Columns 6 and 7 report the coefficients for the belief that democracy works. Finally columns 8 and 9 show the results for the preference of a democratic government to other options. As it was mentioned above, the focus of the analysis is in the measure of uncertainty, σ_{iy}. The table reports that none of the nine measures of preference for democracy are affected by the measure of uncertainty during the impressionable years, except in column 9 where a rather negative weak effect survives at a significant level of 10%. The coefficient estimated for the anomaly variable is statistically not significant in all specifica-

tions. In relation to the control variables, males seem to have a higher appreciation towards democracy. Individuals from larger households are associated with higher preferences for democracy, especially as a type of government. The educational dummies reflect a higher valuation for democracy for higher levels of education. Married individuals have a positive valuation for democracy as a type of government. Having an informal job is both negatively (2002-2006) and positively (2007-2012) associated with the preference for a democratic government. Higher levels of income are systematically associated with higher preferences for democracy in almost all samples. It is just for the question whether democracy works in 2004-2012 that this association turns to be negative.

2.3.2 The Interaction with Autocratic Regimes

In this type of analysis the environment to which cohorts of people are exposed to plays a role in the determination of beliefs. Peruvian history has provided individuals with a variety or regimes to which they may have developed some ideas or beliefs. Anecdotal evidence tells that people from the highlands have a fond memory of Velasco's authoritarian regime, or that some people in cities that were heavily hit by the guerrilla are still fond of Fujimori's government. In consequence, the formation of beliefs about democracy, if related to uncertainty, may have also be shaped by the political context in which uncertainty took place. In a turbulent political period, any political problem may have been unnoticed in the midst of bigger personal problems. To account for the different political settings that nurtured many individuals' experiences, I define a dummy variable to differentiate people who experienced uncertainty only during purely democratic times and individuals that lived under politically problematic and non democratic periods. The goal is to identify individuals whose impressionable years coincided with one of these two definitions. Therefore, I define as non-democratic periods the following years in recent Peruvian history[39]:

1. ***1930-1931***: Under the mandate of general Luis Sánchez-Cerro these two years averaged a Polity index of -5.

[39]I consider only periods in the Twentieth Century, since the oldest Peruvian in the 2002 version of the survey was born in 1905.

2. ***1932-1939***: After the assassination of Sánchez-Cerro, the Constituent Assembly proclaimed Oscar Benavides as president of Peru. Benavides had previously held the position after a revolt in 1914. He was appointed to complete Sánchez-Cerro's mandate. No elections were conducted, and the Polity index for this period was 2. Although it is worth noticing that after the completion of the mandate, democratic elections proclaimed Manuel Prado y Ugarteche as the new president of Peru.

3. ***1948-1956***: The termination of the mandate of Prado y Ugarteche was followed by new democratic elections that proclaimed José Bustamante y Rivero as the next president of Peru in 1945. However, three years later, before the completion of his mandate, general Manuel Odría took office after a Coup d'état, and ruled the country in that position until a brief period between June and July 1950 in which new democratic elections were announced. However, this was a strategy to legitimate his power since he decided to participate in the elections and won, which allowed him to rule the country until 1956. The polity index for these 8 years of Odría mandate averaged -2.

4. ***1962-1963***: After Odría left office in 1956, Prado y Ugarteche was elected president for a second mandate again under a democratic election. His government was assaulted by generals Ricardo Pérez Godoy and Nicolás Lindley in 1962, with the aim to remove the current government and declare new elections. The military junta appointed Pérez Godoy as temporary president of Peru while awaiting for the calling of new elections. After some months, however, general Lindley was aware of general Pérez Godoy's plans to perpetuate in power and revolted against him in March 1963. After removing Pérez Godoy from office, Lindley announced democratic elections in July the same year, in which Fernando Belaúnde was elected president. The average polity index for these two years was -1.

5. ***1968-1980***: Before the termination of Belaúndes's mandate, under allegations of compensating an American oil company, general Juan Velasco launched a military coup that provoked Belaúnde's exile in the United States of America. After 7 years ruling the country, general Velasco experienced a Coup d'état himself. General Francisco Moralez Bermúdez took office in August 1975 and ruled the country until 1980. This was the

year in which he called elections and former president Belaúnde regained power in the polls. The average polity index for the combined period of Velasco and Morales Bermúdez was -5.

6. ***1992-2000***: The second government of Belaúnde witnessed the outbreak of the guerrilla movement, by the end of it. New elections followed in 1985 and Alan García was elected president. García's mandate was characterized by economic disaster that led to an environment of political turmoil by the end of his government. In 1990 new democratic elections were held and Alberto Fujimori was elected president for the period 1990-1995. In April 1992, however, Fujimori shut the Parliament in a move that has been described as a self Coup d'état. After international pressure on Fujimori's decision, he restituted the Parliament by calling to new Parliamentary elections and ruled the country until 2000; year in which he was removed from office after massive protests from all the country. The polity index for this period averaged 1.

Any of the years following within the periods mentioned in the list above is defined as an authoritarian year. I the define the dummy variable *Dem*, which takes the value of 1 if at least one of the impressionable years was not lived under any authoritarian year. I am looking here to isolate the experience of democracy. Therefore, the variable *Dem* takes the value of 0 if all of the impressionable years were lived under an authoritarian regime[40]. Finally, interact this dummy variable with the measures of uncertainty, as shown in equation 2.2.

$$y_{itr} = \alpha_0 + \alpha_1 D^{ro}_{iy} + \beta_1 \sigma^{ro}_{iy} + \beta_2 \sigma^{ro}_{iy} * Dem + Dem + \tau_t + \theta_a + \Gamma X_{it} + v_{it} \tag{2.2}$$

Table 2.3 presents the results of equation 2.2. The interaction of the democratic dummy with the measure of uncertainty, σ_{iy}, seems to unmask nothing: is not statistically significant for any of the measures of democratic preference. Hence, having experienced uncertainty during the impressionable years has no effect on the democratic belief, regardless of the political context uncertainty shared with. Exploring the interaction with the rainfall anomaly yields a similar result. Except for 10% significance in columns 1 and 8, no statistical inference

[40]On average, 86.87% of the individuals fall within the dummy variable. I.e. 13.13% of the individuals lived all their impressionable years under an authoritarian regime.

can be drawn from this table. Uncertainty or rainfall anomaly are not statistically associated to political beliefs even after differentiating by the political context that accompanied the experience of uncertainty or rainfall anomaly.

Table 2.3: Democratic Beliefs and Uncertainty During Impressionable Years and Democracy Exposure

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
D_{iy}	-0.232	0.090	-0.063	0.071	-0.073	-0.113	0.097	-0.175*	0.031
	(0.154)	(0.145)	(0.070)	(0.166)	(0.165)	(0.162)	(0.085)	(0.093)	(0.089)
σ_{iy}	0.044	-0.047	-0.002	-0.010	-0.047	-0.017	0.016	-0.028	-0.023
	(0.064)	(0.064)	(0.031)	(0.069)	(0.075)	(0.072)	(0.038)	(0.041)	(0.039)
$D_{iy}xDem$	0.278*	-0.144	0.081	-0.070	0.071	0.128	-0.086	0.164*	-0.052
	(0.161)	(0.149)	(0.073)	(0.171)	(0.169)	(0.169)	(0.087)	(0.095)	(0.093)
$\sigma_{iy}xDem$	-0.019	0.025	-0.023	0.030	-0.004	0.057	0.035	-0.010	-0.039
	(0.048)	(0.045)	(0.022)	(0.053)	(0.054)	(0.055)	(0.027)	(0.031)	(0.027)
Controls									
Male	0.019**	0.025***	0.035***	0.030***	0.028***	0.015	-0.006	0.031***	0.118***
	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.005)
Household Size	0.003	0.003	0.001	0.005**	0.003	0.009***	-0.000	0.005***	0.005***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Primary	0.057***	0.051***	0.051***	0.058***	0.025**	0.035***	-0.034***	0.080***	0.136***
	(0.010)	(0.010)	(0.004)	(0.010)	(0.011)	(0.011)	(0.005)	(0.006)	(0.006)
Secondary	0.119***	0.098***	0.120***	0.101***	0.089***	0.037**	-0.063***	0.145***	0.187***
	(0.014)	(0.014)	(0.007)	(0.014)	(0.015)	(0.015)	(0.007)	(0.009)	(0.008)
Technical	0.221***	0.195***	0.184***	0.166***	0.189***	0.059**	-0.076***	0.189***	0.249***
	(0.028)	(0.025)	(0.014)	(0.025)	(0.026)	(0.029)	(0.013)	(0.018)	(0.016)
University	0.308***	0.189***	0.212***	0.107***	0.292***	-0.003	-0.086***	0.208***	0.236***
	(0.036)	(0.037)	(0.020)	(0.039)	(0.036)	(0.042)	(0.018)	(0.025)	(0.022)
Married	0.007	0.005	0.002	0.019*	0.011	0.013	-0.008	0.024***	0.043***
	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.005)
Informal Job	0.015	0.002	-0.005	-0.008	0.002	0.002	-0.004	-0.013**	0.013**
	(0.010)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.005)
Unemployed	0.037	0.022	0.007	-0.035	-0.026	-0.004	-0.030*	0.024	0.029
	(0.028)	(0.028)	(0.014)	(0.029)	(0.031)	(0.028)	(0.016)	(0.018)	(0.018)
Household Income	0.019***	0.033***	0.018***	0.020***	0.028***	0.029***	-0.007**	0.023***	0.025***
	(0.007)	(0.007)	(0.003)	(0.008)	(0.007)	(0.008)	(0.003)	(0.004)	(0.003)
R^2	0.362	0.364	0.238	0.365	0.378	0.316	0.298	0.308	0.285
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\gamma_r xYear$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_{ro}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: [1] Standard errors are clustered at the district of birth [2] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

It is possible that the definition of the variable of exposure to democratic regimes, *Dem*, does not capture properly the effect due to the inclusion of the cohort fixed effects. To further investigate this, I instead use a measure of the quality of democracy during the impressionable years that I then interact with the measures of uncertainty. This is explained in section

2.3.4.

Tables 2.4 and 2.5 re-estimate the main empirical equations without the inclusion of the geographical fixed effects. This is just an small exercise to understand the magnitude of the effect being estimated here. The coefficients are higher and show a positive correlation between the measure of uncertainty and the beliefs that democracy works, as reported in table 2.4. There is also a positive association between the belief that democracy is the best type of government and the measure of uncertainty, but only for the period 2002-2006. The belief about importance of democracy is negatively correlated with the uncertainty measure, although only for the 2004-2006 sample.

Table 2.4: Democratic Beliefs and Uncertainty During Impressionable Years: No Geography Fixed Effects

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
D_{iy}	0.070	-0.100**	0.022	0.014	0.192***	0.084	0.142***	0.124***	0.002
	(0.054)	(0.045)	(0.023)	(0.058)	(0.049)	(0.057)	(0.026)	(0.033)	(0.030)
σ_{iy}	0.013	-0.083***	-0.012	0.021	-0.037*	0.062***	0.143***	0.087***	-0.010
	(0.018)	(0.017)	(0.008)	(0.019)	(0.020)	(0.019)	(0.010)	(0.013)	(0.011)
Male	0.036***	0.033***	0.037***	0.023***	0.040***	0.013	-0.001	0.039***	0.119***
	(0.009)	(0.009)	(0.004)	(0.009)	(0.010)	(0.009)	(0.005)	(0.006)	(0.005)
Household Size	0.005***	0.008***	0.002**	0.008***	0.008***	0.011***	0.003***	0.009***	0.004***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Primary	0.047***	0.053***	0.052***	0.065***	0.025**	0.049***	-0.043***	0.091***	0.145***
	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.006)
Secondary	0.113***	0.090***	0.122***	0.113***	0.074***	0.028**	-0.093***	0.154***	0.204***
	(0.013)	(0.013)	(0.006)	(0.013)	(0.013)	(0.013)	(0.007)	(0.009)	(0.008)
Technical	0.190***	0.165***	0.184***	0.158***	0.137***	0.043*	-0.091***	0.216***	0.256***
	(0.025)	(0.024)	(0.013)	(0.021)	(0.024)	(0.025)	(0.012)	(0.016)	(0.015)
University	0.289***	0.149***	0.209***	0.077**	0.191***	-0.039	-0.109***	0.212***	0.243***
	(0.035)	(0.037)	(0.020)	(0.033)	(0.033)	(0.037)	(0.017)	(0.023)	(0.019)
Married	0.005	0.002	0.003	0.013	0.011	0.024**	-0.004	0.034***	0.037***
	(0.009)	(0.009)	(0.004)	(0.009)	(0.010)	(0.009)	(0.005)	(0.006)	(0.005)
Informal Job	0.004	0.005	-0.007*	-0.018*	0.002	-0.015	-0.008	-0.020***	0.005
	(0.009)	(0.009)	(0.004)	(0.009)	(0.010)	(0.009)	(0.005)	(0.006)	(0.005)
Unemployed	0.095***	0.051*	0.012	0.027	-0.003	0.030	-0.029*	0.037**	0.041**
	(0.027)	(0.027)	(0.013)	(0.026)	(0.029)	(0.027)	(0.015)	(0.018)	(0.017)
Household Income	0.043***	0.054***	0.023***	0.061***	0.064***	0.047***	-0.010***	0.036***	0.042***
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	(0.006)	(0.003)	(0.004)	(0.003)
R^2	0.033	0.029	0.046	0.042	0.030	0.040	0.061	0.059	0.093
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	No	No	No	No	No	No	No	No	No
$\gamma_r x Year$	No	No	No	No	No	No	No	No	No
γ_{ro}	No	No	No	No	No	No	No	No	No

Notes: Standard errors clustered at the district of birth; *significant at 10%, **significant at 5%, ***significant at 1%.

Table 2.5: Democratic Beliefs and Uncertainty During Impressionable Years and Democracy Exposure: No Geography Fixed Effects

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
D_{iy}	-0.001	0.352*	0.027	0.196	0.343	0.185	0.395***	0.216*	0.117
	(0.202)	(0.189)	(0.088)	(0.200)	(0.219)	(0.210)	(0.106)	(0.128)	(0.118)
σ_{iy}	0.021	-0.103**	-0.005	-0.003	-0.041	0.035	0.126***	0.095***	0.026
	(0.045)	(0.043)	(0.020)	(0.049)	(0.053)	(0.048)	(0.026)	(0.030)	(0.027)
$D_{iy}xDem$	0.074	-0.475**	-0.005	-0.193	-0.159	-0.105	-0.264**	-0.102	-0.127
	(0.208)	(0.193)	(0.091)	(0.207)	(0.224)	(0.217)	(0.109)	(0.132)	(0.121)
$\sigma_{iy}xDem$	-0.008	0.020	-0.007	0.027	0.004	0.031	0.019	-0.010	-0.042
	(0.049)	(0.047)	(0.022)	(0.053)	(0.057)	(0.053)	(0.028)	(0.033)	(0.029)
Male	0.036***	0.033***	0.037***	0.023**	0.040***	0.013	-0.001	0.039***	0.119***
	(0.009)	(0.009)	(0.004)	(0.009)	(0.010)	(0.009)	(0.005)	(0.006)	(0.005)
Household Size	0.005***	0.008***	0.002**	0.008***	0.008***	0.011***	0.003***	0.009***	0.004***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Primary	0.047***	0.053***	0.052***	0.065***	0.025**	0.049***	-0.043***	0.091***	0.144***
	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.006)
Secondary	0.113***	0.090***	0.122***	0.113***	0.073***	0.028**	-0.093***	0.154***	0.204***
	(0.013)	(0.013)	(0.006)	(0.013)	(0.013)	(0.013)	(0.007)	(0.009)	(0.008)
Technical	0.190***	0.166***	0.183***	0.158***	0.137***	0.043*	-0.092***	0.215***	0.256***
	(0.025)	(0.024)	(0.013)	(0.021)	(0.024)	(0.025)	(0.012)	(0.016)	(0.015)
University	0.289***	0.149***	0.209***	0.077**	0.191***	-0.039	-0.109***	0.211***	0.243***
	(0.035)	(0.037)	(0.020)	(0.033)	(0.033)	(0.037)	(0.017)	(0.023)	(0.019)
Married	0.005	0.002	0.003	0.013	0.011	0.024**	-0.004	0.034***	0.037***
	(0.009)	(0.009)	(0.004)	(0.009)	(0.010)	(0.009)	(0.005)	(0.006)	(0.005)
Informal Job	0.004	0.005	-0.007*	-0.018*	0.002	-0.015	-0.007	-0.020***	0.005
	(0.009)	(0.009)	(0.004)	(0.009)	(0.010)	(0.009)	(0.005)	(0.006)	(0.005)
Unemployed	0.095***	0.051*	0.012	0.027	-0.003	0.030	-0.029*	0.037**	0.041**
	(0.027)	(0.027)	(0.013)	(0.026)	(0.029)	(0.027)	(0.015)	(0.018)	(0.017)
Household Income	0.043***	0.054***	0.023***	0.061***	0.064***	0.047***	-0.010***	0.036***	0.042***
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	(0.006)	(0.003)	(0.004)	(0.003)
R^2	0.033	0.030	0.046	0.042	0.030	0.040	0.061	0.059	0.093
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	No	No	No	No	No	No	No	No	No
$\gamma_r xYear$	No	No	No	No	No	No	No	No	No
γ_{ro}	No	No	No	No	No	No	No	No	No

Notes: [1] Standard errors are clustered at the district of birth [2] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

Results in table 2.5 indicates that the interaction with the dummy variable *Dem* provides little information.

In general, the inclusion of the geographical fixed effects improves the estimation, but the estimation without geographical fixed effects does not provide results that are markedly different from the main results. In any case, the lack of influence of uncertainty on democratic beliefs is robust, as the following sections will show.

2.3.3 Rainfall Shock

Having presented the baseline results, I conduct a series of further explorations to reinforce the previous result: uncertainty is not associated with democratic beliefs. In this section I revisit the definition of anomaly and propose a different measure: a dummy variable to indicate a negative rainfall shock. This redefinition seeks to resemble the measure of macroeconomic shock used by Giuliano and Spilimbergo (2014) and to serve as a further robustness check. The rationale is that rural individuals would find a negative rainfall shock more pervasive than a positive one. I define G_{10} as the measure of negative rainfall shock. It takes the value of 1 if the rainfall experienced in a given year was two standard deviations below the 10 years mean, and 0 otherwise. As with D_{iy}, I match G_{iy} to the district of birth. On average 61.3% of rural individuals have experienced a negative rainfall shock during their impressionable years. In order to keep other things equal, this exercise redefines equations 2.1 and 2.2 as:

$$y_{itr} = \alpha_0 + \alpha_1 G^{ro}_{iy} + \beta_1 \sigma^{ro}_{iy} + \tau_t + \theta_a + \gamma_r + \gamma_{ro} + \gamma_r \tau_t + \Gamma X_{it} + \upsilon_{it} \tag{2.3}$$

$$y_{itr} = \alpha_0 + \alpha_1 G^{ro}_{iy} + \beta_1 \sigma^{ro}_{iy} + \beta_2 \sigma^{ro}_{iy} * Dem + Dem + \tau_t + \theta_a + \Gamma X_{it} + \upsilon_{it} \tag{2.4}$$

Results are shown in table 2.6. Panel a shows the results for equation 2.3 whereas panel b shows the results for the equation with the democracy dummy, 2.4. The inclusion of a measure of negative rainfall shock instead of rainfall anomaly does not change the results. Uncertainty, and negative rainfall shock are not associated with democratic beliefs. In particular, the measure of uncertainty remains statistically insignificant in both models, while

the negative rainfall shock is 10% significant in column 6 of the first panel and in column 8 of the second panel. The interaction of the measure of uncertainty with the democracy dummy is statistically significant at 10% in the second panel of column 3.

Table 2.6: Uncertainty, Negative Rainfall Shocks and Democracy Exposure

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
Panel a. With no democracy interaction									
G_{iy}	-0.005	0.001	0.004	0.000	-0.001	0.018*	0.004	-0.007	0.008
	(0.009)	(0.009)	(0.004)	(0.010)	(0.010)	(0.010)	(0.005)	(0.006)	(0.005)
σ_{iy}	0.020	-0.011	-0.022	0.015	-0.048	0.043	0.045	-0.036	-0.048
	(0.051)	(0.052)	(0.024)	(0.054)	(0.060)	(0.055)	(0.030)	(0.032)	(0.032)
R^2	0.362	0.364	0.238	0.365	0.378	0.316	0.298	0.308	0.285
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Panel b. With democracy interaction									
G_{iy}	0.011	0.012	0.021**	-0.020	0.027	-0.012	0.007	-0.025*	0.004
	(0.021)	(0.021)	(0.010)	(0.022)	(0.025)	(0.022)	(0.012)	(0.013)	(0.013)
σ_{iy}	0.019	-0.028	-0.007	-0.004	-0.053	-0.007	0.021	-0.038	-0.012
	(0.064)	(0.064)	(0.031)	(0.069)	(0.075)	(0.071)	(0.038)	(0.040)	(0.039)
$G_{iy}xDem$	-0.019	-0.013	-0.020*	0.023	-0.033	0.036	-0.003	0.022	0.005
	(0.022)	(0.023)	(0.011)	(0.024)	(0.026)	(0.024)	(0.013)	(0.014)	(0.013)
$\sigma_{iy}xDem$	0.001	0.021	-0.018	0.021	0.006	0.060	0.027	0.004	-0.041
	(0.047)	(0.043)	(0.022)	(0.052)	(0.053)	(0.052)	(0.027)	(0.029)	(0.027)
R^2	0.362	0.364	0.238	0.365	0.379	0.316	0.298	0.308	0.285
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\gamma_r xYear$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_{ro}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: [1] Standard errors clustered at the district of birth. [2] All regressions included controls as in tables 2.2 and 2.3. [3] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

2.3.4 Degree of Authoritarianism

It is possible that having experienced uncertainty during the impressionable years interacted with the democracy dummy may not be capturing properly the information of the cohort. Arguably, the degree of authoritarianism in the decade of 1970 during the two military governments of Velasco and Morales Bermúdez was different from Fujimori's authoritarianism. The Polity index used in figure 2.1 can shed some light in this matter. In this section, I estimate an alternative version of equation 2.2 in which the democratic dummy is replaced by the normalised Polity IV index. The exercise is replicated for the specification that uses the

negative rainfall shock as well (panel b). In concrete, I estimate the two following equations:

$$y_{itr} = \alpha_0 + \alpha_1 D^{ro}_{iy} + \alpha_2 D^{ro}_{iy} * Polity + \beta_1 \sigma^{ro}_{iy} + \beta_2 \sigma^{ro}_{iy} * Polity + \tau_t + \theta_a + \Gamma X_{it} + \upsilon_{it} \quad (2.5)$$

$$y_{itr} = \alpha_0 + \alpha_1 G^{ro}_{iy} + \alpha_2 D^{ro}_{iy} * Polity + \beta_1 \sigma^{ro}_{iy} + \beta_2 \sigma^{ro}_{iy} * Polity + \tau_t + \theta_a + \Gamma X_{it} + \upsilon_{it} \quad (2.6)$$

Table 2.7 presents the results. In the group of beliefs that consider democracy important, no coefficient is statistically significant in the two specifications. The interaction with Polity yields a weak negative association in the third group (column 3) of the first panel. The anomaly coefficient and its interaction are statistically significant at 5% in the first panel, but this results disappears in the two subsequent groups. The belief that democracy is the best type of government despite its imperfections (columns 4 and 5) show no statistical association to uncertainty in any of the two panels. The interaction with Polity, in the first panel, yields a positive association in the first group (column 4) and a negative but significant at 10% for the second group (column 5). In the second panel there is only a positive association in the first group. For the two groups measuring the belief that democracy works (columns 6 and 7) no effect is found for any variable. The two groups of variables for democracy as preferred type of government (columns 8 and 9) are not associated with any measure but with the interaction of uncertainty and Polity in column 9.

In this case, results showed again that a different definition of exposure to authoritarianism does not change the principal result of lack of statistical significance.

2.3.5 Peasants

This section and the next one revisit the definition of rural individuals. The standard definition of rural household in ENAHO implies people living in scattered villages and/or agglomerations smaller than 400 households. This proportion is still important in Peru: in 2012, 25% of the Peruvian households were rural.

However, individuals could be living in a rural setting but the type of activity or occupation they are engaged to may be very different from farming, which does not make them vulnerable to swifts in weather conditions. Figure 2.2 shows that the proportion of indi-

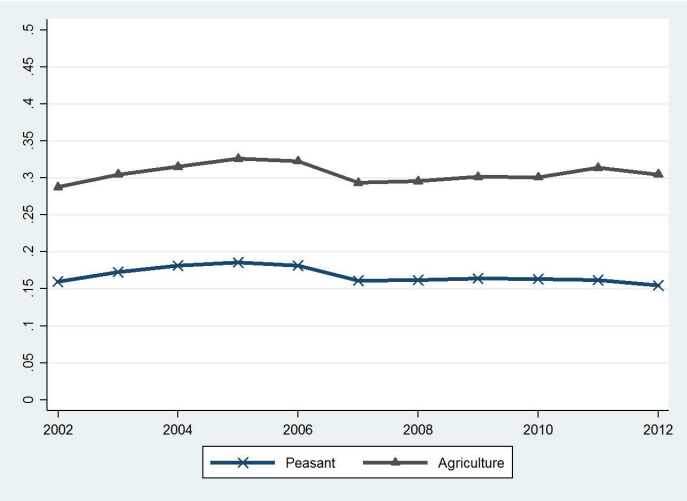
Table 2.7: Uncertainty And Shocks Interacted with Polity Index

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
Panel a. Using Anomaly									
D_{iy}	-0.355**	0.035	-0.015	0.227	-0.355*	0.019	0.056	-0.115	0.096
	(0.173)	(0.164)	(0.081)	(0.182)	(0.191)	(0.183)	(0.094)	(0.101)	(0.101)
σ_{iy}	0.110	-0.019	0.018	-0.124	0.053	-0.073	0.028	0.014	-0.002
	(0.077)	(0.074)	(0.033)	(0.081)	(0.087)	(0.086)	(0.039)	(0.045)	(0.042)
$D_{iy}xPolity$	0.488**	-0.098	0.027	-0.279	0.414*	-0.014	-0.049	0.112	-0.146
	(0.215)	(0.195)	(0.098)	(0.228)	(0.225)	(0.234)	(0.111)	(0.122)	(0.124)
$\sigma_{iy}xPolity$	-0.119	-0.009	-0.055*	0.200**	-0.154*	0.146	0.025	-0.075	-0.073*
	(0.083)	(0.075)	(0.032)	(0.088)	(0.090)	(0.092)	(0.034)	(0.046)	(0.040)
R^2	0.363	0.364	0.238	0.365	0.379	0.316	0.298	0.308	0.285
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Panel b. Using Rainfall Shock									
G_{iy}	0.015	0.022	0.011	-0.023	0.073**	-0.030	0.022	-0.021	0.001
	(0.029)	(0.029)	(0.012)	(0.030)	(0.032)	(0.031)	(0.015)	(0.018)	(0.015)
σ_{iy}	0.074	-0.015	0.017	-0.104	0.003	-0.033	0.018	0.012	0.016
	(0.077)	(0.074)	(0.034)	(0.082)	(0.088)	(0.087)	(0.041)	(0.046)	(0.043)
$G_{iy}xPolity$	-0.031	-0.032	-0.010	0.035	-0.109**	0.074	-0.027	0.023	0.011
	(0.043)	(0.042)	(0.018)	(0.044)	(0.046)	(0.046)	(0.021)	(0.026)	(0.022)
$\sigma_{iy}xPolity$	-0.072	0.010	-0.050	0.161*	-0.055	0.098	0.039	-0.072	-0.088**
	(0.083)	(0.076)	(0.033)	(0.089)	(0.090)	(0.092)	(0.038)	(0.048)	(0.042)
R^2	0.362	0.364	0.238	0.365	0.379	0.316	0.298	0.308	0.285
Observations	16394	16456	42289	15433	15069	16552	56385	40145	52165
Clusters	12540	12227	24541	11878	11345	12624	30902	24091	28378
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\gamma_r xYear$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_{ro}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: [1] Standard errors clustered at the district of birth. [2] All regressions included controls as in tables 2.2 and 2.3. [3] Polity Index is normalised to the max and min of the complete sample for all countries. [4] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

viduals that were peasants in the sample fluctuates around 16% whereas the proportion of individuals working in an agricultural activity fluctuates around 30 %.

Figure 2.2: Proportion of Individuals Declared as Peasants or Engaged in Agricultural Activities: 2002-2012



Notes: [1] Data source: ENAHO 2002-2012. [2] Unweighted averages. [3] Peasant is the proportion of individuals who declared their main occupation is farming. Agriculture refers to the proportion of individuals whose main labor activity was classified in chapters 1,2 and 5 of the ISIC Rev. 3.

In this section I constrain the sample to those individuals that declare their main occupation is being a farmer (peasant). This reduces the sample size by 75%. I then re-estimate equation 2.1 considering rainfall anomaly (panel a) and rainfall negative shock (panel b). Results are shown in table 2.8.

In general results are statistically the same as in the previous cases: no effect of uncertainty on democratic beliefs (except for a effect statistically significant at 10% in column 8 of panel a). More surprisingly is that the negative rainfall shock modeled in panel b has no statistically significant result except in the last column.

2.3.6 Agriculture Workers

This section uses the agriculture classification outlined above and restricts the sample to those individuals who declared that their main income activity in related to agriculture as defined by the International Standard Industrial Classification of All Economic Activities (ISIC), Rev.3[41]. Results are reported in table 2.9

[41]Chapters 1, 2 and 5

Table 2.8: Democratic Beliefs and Uncertainty During Impressionable Years and Democracy Exposure: Peasants

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
Panel a. Using Anomaly									
D_{iy}	0.089	-0.140*	-0.008	-0.101	-0.006	-0.074	-0.008	-0.026	-0.006
	(0.103)	(0.083)	(0.039)	(0.112)	(0.100)	(0.119)	(0.043)	(0.049)	(0.047)
σ_{iy}	-0.022	-0.022	-0.027	-0.131	-0.015	0.075	0.059	-0.121*	0.029
	(0.124)	(0.138)	(0.053)	(0.139)	(0.158)	(0.139)	(0.070)	(0.071)	(0.066)
R^2	0.564	0.546	0.460	0.598	0.587	0.548	0.503	0.469	0.460
Observations	5860	5825	14240	5474	5269	5939	19179	14756	18192
Clusters	5148	4941	10540	4818	4501	5214	13755	11124	12772
Panel b. Using Rainfall Shock									
G_{iy}	-0.029	0.004	0.009	0.003	-0.003	0.023	0.004	0.009	0.031***
	(0.021)	(0.022)	(0.009)	(0.023)	(0.025)	(0.023)	(0.011)	(0.012)	(0.011)
σ_{iy}	-0.060	0.039	-0.019	-0.112	-0.014	0.105	0.063	-0.110	0.049
	(0.124)	(0.135)	(0.053)	(0.139)	(0.154)	(0.138)	(0.069)	(0.070)	(0.065)
R^2	0.564	0.545	0.460	0.598	0.587	0.548	0.503	0.469	0.460
Observations	5860	5825	14240	5474	5269	5939	19179	14756	18192
Clusters	5148	4941	10540	4818	4501	5214	13755	11124	12772
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\gamma_r x Year$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_{ro}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: [1] Standard errors clustered at the district of birth. [2] All regressions included controls as in tables 2.2 and 2.3. [3] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

Table 2.9: Democratic Beliefs and Uncertainty During Impressionable Years and Democracy Exposure: Agricultural Workers

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
Panel a. Using Anomaly									
D_{iy}	0.055	-0.069	-0.004	-0.021	-0.041	0.052	0.016	-0.007	0.017
	(0.056)	(0.051)	(0.023)	(0.062)	(0.058)	(0.061)	(0.026)	(0.031)	(0.031)
σ_{iy}	0.023	-0.048	-0.003	0.021	-0.088	0.082	0.079**	-0.069*	-0.037
	(0.060)	(0.064)	(0.028)	(0.066)	(0.073)	(0.066)	(0.035)	(0.038)	(0.036)
R^2	0.403	0.389	0.267	0.405	0.420	0.372	0.333	0.335	0.308
Observations	13793	13753	36276	12913	12533	13943	47991	33855	44831
Clusters	11053	10697	22222	10402	9879	11161	27987	21691	25696
Panel b. Using Rainfall Shock									
G_{iy}	-0.017	0.004	0.004	0.006	-0.000	0.020*	0.004	-0.008	0.008
	(0.010)	(0.011)	(0.005)	(0.011)	(0.012)	(0.011)	(0.006)	(0.006)	(0.006)
σ_{iy}	-0.002	-0.024	0.001	0.030	-0.075	0.086	0.078**	-0.073*	-0.034
	(0.059)	(0.063)	(0.028)	(0.066)	(0.071)	(0.065)	(0.034)	(0.037)	(0.035)
R^2	0.403	0.389	0.267	0.405	0.420	0.372	0.333	0.335	0.308
Observations	13793	13753	36276	12913	12533	13943	47991	33855	44831
Clusters	11053	10697	22222	10402	9879	11161	27987	21691	25696
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_r	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\gamma_r xYear$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_{ro}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: [1] Standard errors clustered at the district of birth. [2] All regressions included controls as in tables 2.2 and 2.3. [3] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

In line with previous results, uncertainty during the impressionable years has no statistically significant effect on democratic beliefs, except in column 7 in panels a and b. Column 8 shows a weak negative effect. The rainfall anomaly has no effect at all in any of the measures of democratic preferences. In panel b, just column 6 shows a positive association between a negative rainfall shock and the belief that democracy is the best option despite its imperfections, but it is statistically significant at 10%

2.3.7 Current Uncertainty

After having examined the possible effect of uncertainty during the impressionable years and finding no statistically significant and conclusive result, this section explores the possibility that current uncertainty has some effect on the determination of democratic beliefs. In order to do this I use the measure of current uncertainty explained in section2.2.2. Equation 2.1 takes the following form:

$$y_{itr} = \alpha_0 + \alpha_1 D^r_{10} + \beta_1 \sigma^r_{10} + \tau_t + \theta_a + \gamma_{ro} + \gamma_{ro}\tau_t + \Gamma X_{it} + v_{it} \tag{2.7}$$

The definition of D_{10} and σ_{10} prevent the use of fixed effects for the current district: unlike all previous estimations, both variables are matched with the current district and year combination. This may impede to control for current political trends developing in the current district, that may shape the democratic preference of the individual. This estimation controls for trends in the district of birth of the individual, and I keep the same individual and household characteristics as in 2.1. Table 2.10 shows the results.

Table 2.10: Democratic Beliefs and Current Uncertainty

	Is Important			Is Best		Works		Government	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	02-04	04-06	07-12	02-04	04-06	02-04	04-12	02-06	07-12
Panel a. Using Anomaly									
D_{10}	0.012	0.016	-0.006	0.057	0.178***	0.100**	0.044**	0.053**	0.038**
	(0.045)	(0.036)	(0.014)	(0.041)	(0.039)	(0.042)	(0.018)	(0.023)	(0.018)
σ_{10}	-0.084**	-0.101***	0.021	0.036	0.062	0.056	0.083***	0.047*	0.030
	(0.042)	(0.039)	(0.019)	(0.044)	(0.042)	(0.047)	(0.022)	(0.027)	(0.024)
R^2	0.315	0.311	0.242	0.318	0.330	0.278	0.292	0.285	0.282
Observations	16239	16259	41845	15282	14880	16401	55753	39767	51689
Clusters	12458	12145	24395	11795	11266	12544	30768	23986	28237
Using Rainfall Shock									
G_{10}	-0.047**	0.061***	-0.002	-0.028	0.039*	-0.022	-0.006	0.001	0.019
	(0.021)	(0.019)	(0.010)	(0.021)	(0.020)	(0.022)	(0.011)	(0.012)	(0.013)
σ_{10}	-0.078*	-0.106***	0.022	0.038	0.051	0.055	0.080***	0.044	0.028
	(0.042)	(0.039)	(0.019)	(0.044)	(0.042)	(0.047)	(0.022)	(0.027)	(0.024)
R^2	0.316	0.312	0.242	0.318	0.329	0.277	0.292	0.285	0.282
Observations	16239	16259	41845	15282	14880	16401	55753	39767	51689
Clusters	12458	12145	24395	11795	11266	12544	30768	23986	28237
Fixed Effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\gamma_{r0}xYear$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
γ_{ro}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: [1] Standard errors clustered at the district of birth. [2] All regressions included controls as in tables 2.2 and 2.3. [3] Coefficients that are statistically significant are denoted by the following system: *** 1%, ** 5%, and * 10%.

If there is any statistical effect, that is for the group of democratic belief related to how important democracy is. In both panels current uncertainty is negatively associated with the belief that democracy is important. Only in panel b a negative rainfall shock is also negatively associated with the belief that democracy is important. However, column 3, which belongs into that group of variables shows no effect of uncertainty on the dependent variable. If democracy is the best option despite its imperfections has no relation to current uncertainty in both panels. Just in column 5 a positive association emerges for both, the anomaly and the negative rainfall shock. For the groups of variables that measure the belief that democracy works, only for the second group (column 7) the measure of uncertainty shows a positive link with the belief that democracy works. The negative rainfall shock has no effect on this kind of belief, but the anomaly does: it is positively associated. Finally, for the group of variables that measure the preference of a democratic government, just the anomaly shows a positive association. Uncertainty is weakly associated in column 8.

The main conclusion from this section is that although some association between current

uncertainty and democratic beliefs emerges, it is not conclusive[42].

2.4 Allowing for different Impressionable Age

This section presents a substantial robustness check. Although in this chapter I have followed the definition of impressionable years as proposed by Giuliano and Spilimbergo (2014), it is possible to believe that the impressionable years age could not be restricted to that period.

In this exercise I completely relax that assumption and instead of considering impressionable years the period between 18 and 25. Instead, I use the uncertainty and anomaly experienced by the individual at every year of their life, separately. The advantage of this procedure is twofold. First, as explained above, this allows me to consider different formative period in the life of the individual. Second, although I find no effect of uncertainty on the determination of democratic belief, this is a falsification test: individuals experiencing uncertainty at age 0, 1 or 2 should not have any different preference than others.

In particular, I re-run equation 2.1 independently for each year in the range of 0-60[43]. Figures 2.3 to 2.10 present the results. In all cases I report the estimates of D[44] (left panel) and σ (right panel) for the individual year. The vertical axis reports the estimated coefficient whereas the horizontal axis reports the age. I also report the confidence interval of each coefficient, and the the t-statistic when significant (on the right axis of each panel).

2.4.1 Democracy is Important

I start with the group of variables on the importance of democracy. For the first period, 2002-2004 (figure 2.3) the effect of rainfall anomaly at any age of the individual does not seem to follow any pattern with aging and the coefficient is not statistically significant at any age except when the individual is 6 years old. At older ages the estimation of the coefficient is very inaccurate due to the small number of individuals included in the estimations. The

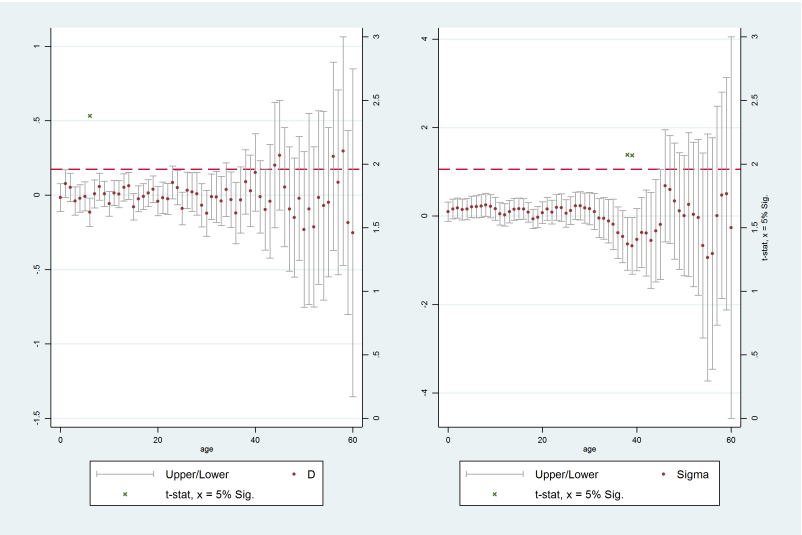
[42]For instance, if I include current district fixed effect, the estimated associations vanish for all variables. The inclusion of the current district fixed effect is possible because in practice I am matching current uncertainty depending on the month of the survey.

[43]The exercise involved estimations until age 65, but the estimated standard errors in that later group were too large that impeded a proper visualization of the results.

[44]I replicated the exercise using the rainfall anomaly, G instead of D but results were the same. Figures available upon request.

effect on uncertainty (right panel) reaches its minimum when the individual is around 38-39 and is statistically significant only for those two years.

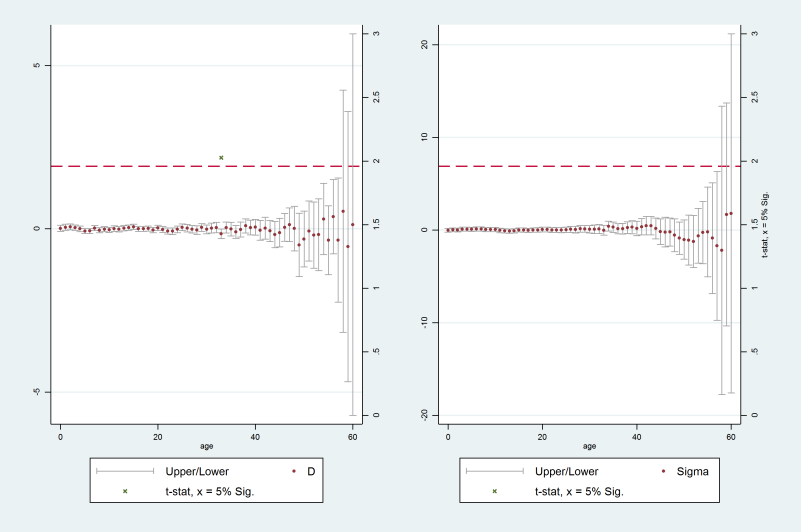
Figure 2.3: Democracy is Important: 2002-2004



Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

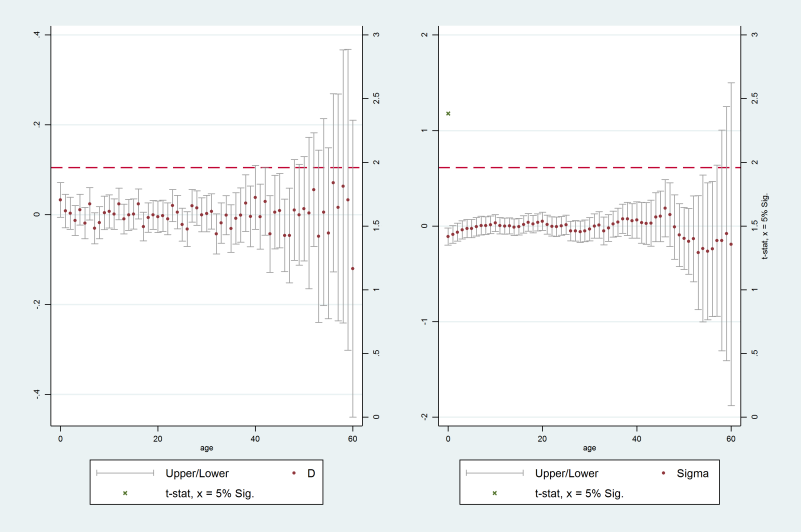
However, the estimates of D and σ in the next two periods, 2004-2006 (figure 2.4) and 2007-2012 (figure 2.5) are all not statistically significant except for one year: the estimate of D in 2004-2006 is statistically significant when the individual is 33 years old, which has no relation with the previous result of an statistically significant effect of anomaly when the individual was 6 years old. There are only 4 years of difference at much between one survey group and the other, too short to argue that it is the same effect found in the cohort of people that experienced a significant effect if they were exposed to rainfall anomaly at an age of 6.

Figure 2.4: Democracy is Important: 2004-2006



Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

Figure 2.5: Democracy is Important: 2007-2012

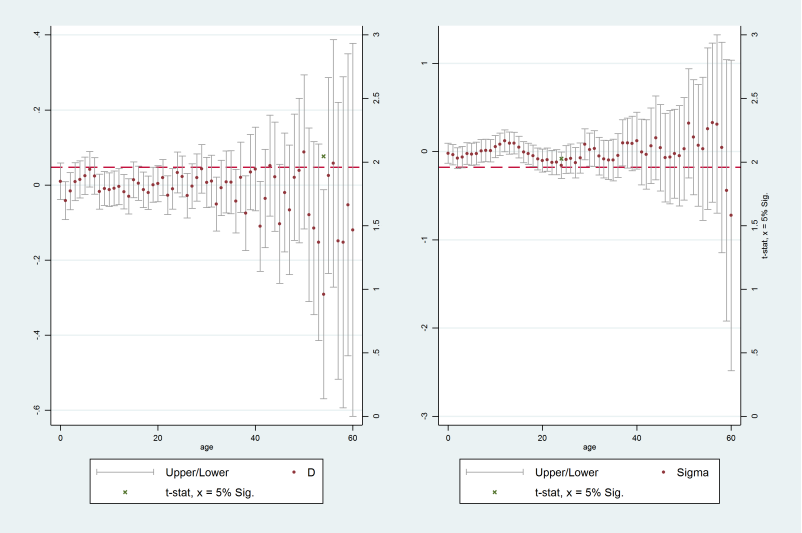


Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

2.4.2 Democratic Government is Preferred

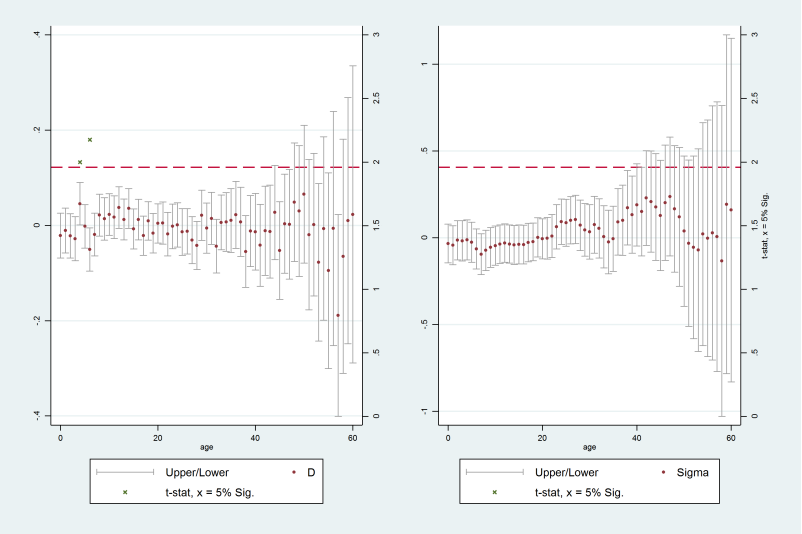
This section examines whether rainfall uncertainty and anomaly have any effect on the belief that democracy is a better type of government in contraposition to an authoritarian or military regime. As in the case of the model for the importance of democracy, the estimation of the anomaly, D, seems not to follow a clear patter along age. In both figures (2.6 and 2.7) the left panel shows an estimate that is imprecisely estimated around zero and becomes more imprecise with age.

Figure 2.6: Democracy is the Preferred Type of Government: 2002-2006



Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

Figure 2.7: Democracy is the Preferred Type of Government: 2007-2012



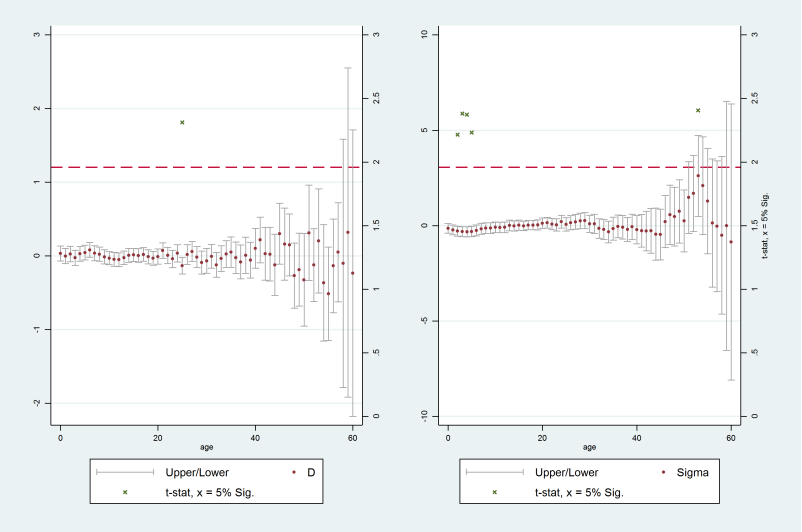
Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

The estimation of the coefficient of the measure of uncertainty, σ, is not statistically significant in any year. In figure 2.6 it reaches a minimum when the individual is aged around 24-30, whereas in figure 2.7 the lowest estimated coefficient takes place around the age of 9-10. In both periods, the estimation of the coefficient becomes very imprecise for later ages.

2.4.3 Democracy Works

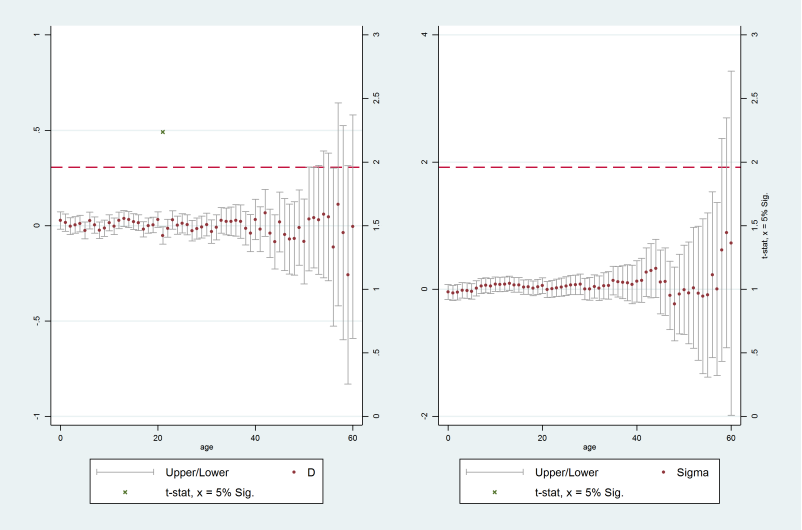
In this section the dependent variable is the belief that democracy is working. In terms of the anomaly (left panel in figures 2.8 and 2.9), this coefficient is not statistically significant through the age of the individual, except for anomaly at age 21 in figure 2.8 and age 24 in figure 2.9.

Figure 2.8: Democracy Works: 2002-2004



Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

Figure 2.9: Democracy Works: 2004-2012



Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

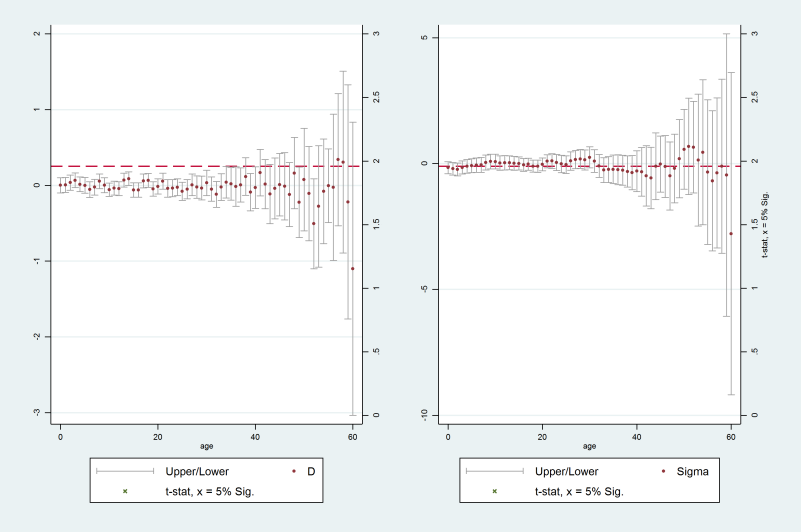
The measure of uncertainty reaches its largest negative effect at early stages of life, and is statistically significant for ages 3, 4, 5 and 6 in figure 2.8. However, in the second sample

in figure 2.9 such effect disappears. As in the previous cases, the estimation of coefficient at late ages becomes very imprecise.

2.4.4 Democracy is Best

Finally, the las variable is the belief that democracy despite its imperfections is the best type of government. In this case I can only report the results for the 2007-2012 (figure 2.10) sample due to constraints with the degrees of freedom. The inclusion of fixed effects of current and birth district and the year district leaves few degrees of freedom for the estimation in the first sample (2002-2004) specially at later ages[45].

Figure 2.10: Democracy is the Best Option: 2002-2004



Notes: [1] Left axis is the estimated beta coefficient for individual ages ranging from 0 to 60. [2] Left panel is the estimate for D_{iy}; right panel is the estimate for σ_{iy}. [3] Confidence intervals at 95%. [4] Red line is the critical value at 5% significance level, 1.96. [5] T-stat shown only when larger than 1.96.

Along different age years, none of the two coefficients are statistically significant. Therefore experiencing rainfall anomaly or rainfall uncertainty at any age does not have any effect on the belief that democracy, in spite of its imperfections, is the best type of government.

[45]The partial graph (not reported), estimated up to age 45, however, shows similar results as in the previous cases.

2.5 Conclusions

In this chapter I have analyzed the relationship between economic uncertainty and democratic beliefs. In particular, I explored whether such relationship takes place during the impressionable years of the individual; and whether it also relates to the exposure to different political regimes. I find no statistically significant relationship to this respect, and the result is robust to different specifications. Current uncertainty plays no role in the determination of democratic beliefs either.

Democratic beliefs (measured in four dimensions: importance of democracy, performance, best option and type of government) of rural individuals from Peru do not respond to economic uncertainty during the impressionable years (18-25); and the exposure to uncertainty during purely democratic or authoritarian regimes adds no explanation to this lack of association. Rainfall anomaly and rainfall negative shock do not affect democratic beliefs either.

On further robustness checks, uncertainty combined with a measure of degree of authoritarianism do not have an statistically significant effect on the determination of democratic beliefs either. The measures of rainfall anomaly and negative rainfall shock combined with the degree of authoritarianism do not show association with the democratic beliefs.

A narrower definition of rural individuals shows that this lack of association between economic uncertainty and democratic beliefs remain. Individuals working in agriculture of self classified as peasants do not define their democratic beliefs due to uncertainty or rainfall anomalies. Current uncertainty have statistically significant effect in two of the nine measures of democratic beliefs: current uncertainty is negatively associated with the belief that democracy is important during the periods 2002-2004 and 2004-2006. However this result is not statistically significant for the period 2007-2012, and more importantly, it may be the consequence of not controlling properly for current trends in the geographical location of the individuals.

Allowing for different formative years does not change the main result. Experiencing uncertainty at any moment in life has no effect on the formation of democratic beliefs.

The main conclusion from this chapter, therefore, is that economic uncertainty plays no

role in the determination of democratic beliefs. This may be because risk exposure does not motivates individuals to pursue a more cooperative behavior that transmigrates into a better democratic perception. Or it may be the case that democratic beliefs have a different type of determinants. An earlier version of this paper also evaluated whether conflict had any effect, using the same empirical strategy as in chapter 2 to no avail.

Further explorations may take into consideration other types of uncertainty or shocks experienced by the individuals in the determination of democratic beliefs.

3 Long Run Empirical Regularities of Mining Booms at the Local Labor Market Level

3.1 Introduction

Over the past decade commodity prices boomed to the benefit of many countries endowed with natural resources. When a producing country experiences a price boom, revenues increase from a macro perspective and GDP also booms, mainly fueled by higher revenues. However, from a micro perspective, it is not clear who gains from the mining boom or what are the linkages with other industries. Large scale mining is generally regarded as an industry with little contribution to the local labor market due to its high dependency on capital. This concern is of course, particularly relevant in a country whose main activity is mining.

The goal of this chapter is to provide an understanding of the degree of connection between the natural resources sector and other industries. The way I analyze this connection is through the local labor markets.

This chapter, therefore, merges two streams of literature: one that understands the interactions within the local labor markets and one that is permeated with the idea of the "Natural Resource Curse", or "Dutch Disease". The current emergence of the literature on local labor markets has prompted a set of new tools that take the city or any other small geographical unit as the piece of observation, and naturally links the response of labor variables to local demand or productivity shocks.

In particular, the literature on local labor markets tries to understand the local effects of a local demand shock on employment, population and wages (Moretti; 2010, Notowidigdo; 2013), and its interaction with any policy intervention (Kline and Moretti; 2014). Moretti (2011) offers a coherent framework based on the works of Roback (1982), Topel (1986), Bartik (1991) and Blanchard and Katz (1992) that motivates further developments. In particular, the author provides the theoretical framework to understand the effects of a productivity shock in local markets. In his formulation, the effect come as a consequence of the interaction of labor and housing markets in a context where individuals take decisions based on the wage differences across cities. Empirically, Moretti (2010) proposes an estimation strategy based

on the Bartik type of shocks[46] to understand the creation of jobs in the non-traded sector as a consequence of a demand shock in the traded sector. Notowidigdo (2013) develops a theoretical model to explain the effect of local demand shocks conditional on the degree of mobility attributed to different types of labor. Theoretically, the author contributes to the discussion by including two types of workers: high skilled and low skilled. His estimates suggest that low skilled workers react less to local demand shocks. Kline and Moretti (2014) also develop a theoretical model to understand the impact of local based interventions. One extension of their theoretical model relates the effect of productivity shocks to agglomeration.

More recently, Alcott and Keniston (2015) develop a theoretical model to evaluate the local impact of a shock in the natural resource industry given its links to intermediate industries and local firms. Their model is tested with data from the United States and assumes, however, only one type of labor. The theoretical framework developed in this chapter draws inspiration from this study.[47]

On the other hand, the literature on the "Natural Resource Curse" or "Dutch Disease"[48] has traditionally taken a macro perspective. At the macro level, the effects of a Dutch Disease can be listed as: (i) real exchange rate appreciation, (ii) fall in manufacture production (iii) fall in the profits for the traded manufacturers. The availability of data at a smaller geographical level allows to re-visit this literature through the lenses of the local labor markets[49]. Beyond the question whether the abundance of natural resources is good or not for long-run growth, the question about the mechanisms operating within the country is equally relevant. What are the local effects of a boom in oil or mineral production? Recent empirical studies at the micro level have attempted answers to these questions. Black et al. (2005) evaluate the impact the coal mining boom, peak and bust for a group of producing cities in the United States. This chapter draws empirical inspiration from this work, but extends the analysis to

[46]The original idea belongs to Bartik (1991). A Bartik instrument exploits the differences in the industry composition across cities compared to the national industry composition. The source of the identification comes from the interaction of the national growth rate of the industries with the industry shares at the city level. A notable application for the housing markets in the U.S. can be found in Guerrieri et al. (2013).

[47]Which, is also related to the preliminary discussion about the interaction between tradable and non-tradable industries presented in Moretti and Thulin (2013)

[48]Ideas championed by Jeffrey Sachs and Andrew Warner in the nineties (Sachs and Warner; 1995, Sachs and Warner; 1999)

[49]Although the effect on real exchange rate seems inherently embedded in the macro perspective, it is easily implemented at the local economy, since the real exchange rate can be expressed as the ratio between the prices for traded goods and non-traded goods.

the national level and by type of labor.

In the developing country setting, much effort has been placed on identifying short-term impacts, but there is a lack of evidence on the long-term impact or the heterogeneity of the response by the type of worker, as well as a better documentation of the channels. Aragón and Rud (2013), another work from which I draw empirical inspiration, explore the backward linkages of a demand shock from a gold mine. The authors build upon the framework in Moretti (2010) to explore local multipliers and find that the mining local purchases impact positively local income. In this chapter I take a finer look at the labor markets and focus on wages instead of household income.

Fafchamps et al. (2015) propose a broad empirical model to understand the booming gold activity in Ghana. Using night lights data, the authors conclude that mining shocks can predict proto-urbanization in the area surrounding the mine. Interestingly, they also find that the results are not reversed once the mine is closed. Loayza et al. (2013) studies the effect of mining on poverty and inequality at the local level in Peru, and find that an increase in the local revenues derived from the mining activity is associated with lower levels of poverty but higher consumption inequality.

This chapter seeks to contribute by proposing an empirical strategy to address the main implications related to the effect of mining booms at the local level. I use a sample of 1043 districts from Peruvian highlands that were close to a mine in 1993 and observe their employment and population response following a boom in the mineral prices from 1993 to 2007. Other things equal, districts close to the mine and those in the surrounding area can be comparable. The challenge consist, precisely, on gathering the appropriate control group for comparison.

As a preview, results suggest that large-scale mining activity boosted total employment in the local economy, which is constituted by the set of districts within a 100km distance to the mine. Both, high skilled and low skilled employment rates grew as a result of this. Wages in the mining area increased, and such increase was focused in agricultural earnings. This result provides enough evidence to conclude that mining booms affect local agricultural economies through the *spending effect* as outlined by Corden and Neary (1982). The point estimates for the effect on low skilled wages are higher than the point estimates for the effect

on high skilled wages, but they are not statistically different. Finally, the faster increase in the employment rate was explained by a faster increase in the employment rate of locals rather than migrants. This effect is also heterogeneous by industry: high skilled locals fill the new mining employment opportunities, while low skilled locals explain the increase in agricultural employment.

For the remainder, this chapter is organized as follows: section 3.2 defines the theoretical framework to guide the analysis. Section3.3 details the data used in the analysis, explains the selection of the districts and specifies the identification strategy. Section 3.4 presents the results. Finally section 3.6 concludes.

3.2 Conceptual Framework

The theoretical basis for the analysis is grounded on the local labor markets literature (Moretti; 2011). In this section I develop a narrative to guide the empirical results. The idea behind is the existence of heterogeneous firms that use the same technology but respond differently to productivity shocks in the natural resources industry, and demand different types of labor.

The unit of analysis is the district or locality where different types of firms operate. Firms may belong to different groups: natural resources, tradable and non-tradable. Natural resource firms sell their output in the international markets for mineral commodities. Firms in the tradable sector sell in the national, local or international markets[50]. Firms that operate in the local economy, i.e. firms in the non-tradable sector sell locally. Non-tradable sector is comprised by all other firms with the exception of government. These three types of firms compete for two types of workers in the locality: high skilled workers (H_i) and low skilled workers (L_i), who receive different wages: w_H and w_L. The national population, Pop, is fixed, so changes in the district population come in the form of people moving from one city to the other depending on the relative city wages. The national stock of high skilled H and low skilled L workers is also fixed.

In a setting like this, it is possible to think of workers, who supply labor, maximize their utility depending on aggregate wages for their type, amenities and a their location prefer-

[50]Tradable sector could be constituted by manufacture and agriculture firms.

ence (also depending on their type).[51] Consequently, the location preference can also be understood as mobility cost[52]. The idea is simple, high skilled individuals may exhibit less attachment to cities and therefore their location preference is less rigid, less strong, which ultimately can also be understood as lower mobility costs.[53]

Labor demand, on the other hand is determined by firms. Firms demand labor locally through the maximization of their profits. Each sector combines capital and a labor, where labor is combined assuming some degree of substitution between high and low skilled workers.

It is also possible to think that all industries are affected by productivity shocks in the natural resource sector. As in Alcott and Keniston (2015), the natural resources sector impacts other industries via the change in revenues. An increase in local revenues derived from the boom in the natural resources sector can be interpreted as a demand shock for the the industry selling locally. To the degree that the local manufacture industry produces intermediate good used as inputs by the natural resources sector, the boom in the mining sector also creates a demand shock for the manufacture producers. Moretti and Thulin (2013) provide another useful theoretical discussion which highlights the connections among sectors. In particular, the authors highlight the relevance of migration costs in the determination of wages.

Empirically I will evaluate the labor demand shock as a ***price shock*** in the natural resource sector: a productivity shock in the natural resources sector represents in income shock for workers in that industry, who happen to live locally. The increase in their rents spills over the local economy through consumption of local goods, which increases prices and the rents of the non-tradable or service sector. The same explanation is less clear for the tradable sector. Its tradable nature implies that the prices they receive do not change, and on the contrary, the costs they face may increase, if the wages of the workers they employ increase, which will represent a negative shock. However, it is also possible that there are linkages between the natural resources sector and the tradable, which ultimately would imply a positive spillover.

[51] For instance, Moretti (2011) differentiates the utility derived from amenities by each type of labor.

[52] As in Notowidigdo (2013)

[53] Notowidigdo (2013) provides an insightful discussion about the location preferences. High preferences for one location, or high migration costs, imply a less responsive labor supply. If location preferences are irrelevant or there are no mobility costs, labor supply is very sensitive to wage differences between cities.

Equilibrium in the labor market takes place when labor demand and supply intersect. If the mobility cost of high skilled workers is smaller than the mobility cost of low skilled workers and for a given elasticity of substitution between high and low skilled workers, employment in the high skilled sector should increase faster than employment in the low skilled segment of workers following a productivity shock in the natural resource sector.

The response of wages to the shock in the natural resource sector depends on the location preferences by each type of worker and the elasticity of substitution between high and low skilled workers. In the extreme case that both types of workers have the same location parameters or mobility costs, their wage response is identical, regardless of the elasticity of substitution. Depending on the differences in the location preferences, the impact on wages varies.

If high skilled workers are more mobile than low skilled workers, the initial push on wages, is offset by the differential labor supply response, and wages for the low skilled workers should increase. However, this balancing force is attenuated by the elasticity of substitution of labor: if there is perfect substitution, the lack of mobility of low skilled workers could be compensated by high skilled workers taking up low skilled jobs which now are better paid.

3.3 Data and Identification Strategy

The empirical analysis uses 1043 districts from the highlands in Peru that I observe between 1993 and 2007. These 1043 districts were selected among the 1791 available that constitute the total number of Peruvian districts in 1994. The basic criteria for the selection sought to create a comparable set of districts in 1993, before the boom in the prices of minerals. All districts are located in the highlands of Peru, above 1,800 meters above the sea level. No large-scale mining activity is reported below 2,000 meters above the sea level. Hence, the selection process excludes relatively wealthy districts from the coast, and the relatively isolated districts from the jungle.

For the analysis I construct information at the district level about employment, population and migration by industries and skills. While for the analysis on wages I use individual level data for a sample of people living in those 1043 districts.

To estimate the impact of the boom in the mining activity, I create a measure of the boom in mineral prices. In concrete, I use information of 27 mines that were active in 1993-1994 in Peru and define a distance threshold for their influence: 100km. Every district's capital[54] whose Euclidean distance to any of the 27 mines in 1993 was smaller or equal than 100km was considered a mining district. Districts beyond that distance threshold but no more than 200km are selected as *control* group.

The unit of analysis for the variables on employment, population and migration, is the district. For wages, the unit of analysis is the individual living either in the mining or control districts. The variable of interest, the boom in the mineral prices, is measured as a compound of mines, production and prices at the district level (explained in section 3.3.3).

In this section I detail the data I use, as well as the empirical strategy to identify the effect of the boom in mineral prices, which led to the selection of these 1043 districts.

3.3.1 Data

For employment I use the population censuses of 1993 and 2007 in the estimations, and the population censuses of 1993 and 1981 to check the comparability of the districts in 1993. To estimate the effect of the boom in the mineral prices on wages I utilize data contained in the employment module from the National Household Survey (ENAHO in Spanish). This survey is collected annually from 1997 by the National Institute of Statistics of Peru. Data about mine location and production are available from InterraRMD (2013), which provides the GPS location of all large-scale mines in Peru, as well as physical production from the beginning of the 20th century in some cases. I discuss all of them in detail.

Mineral Data

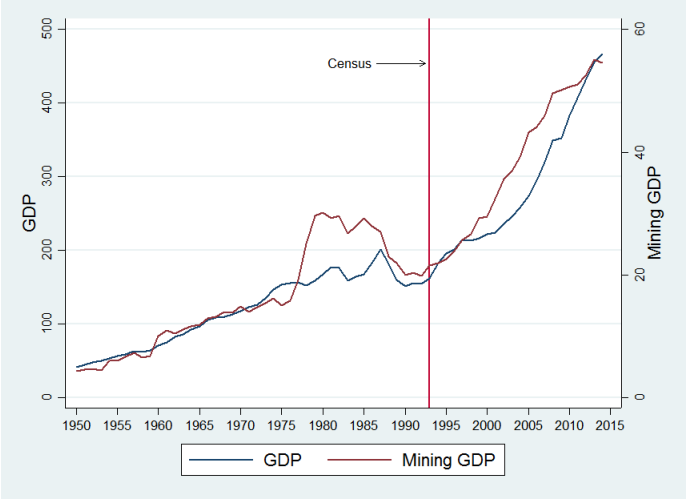
Data about production and geographic location of mines are available from the Raw Material Data (RMD) from InterraRMD (2013). This dataset has records for 633 mine projects through the period of 1900-2014. Each mine has a GPS location that I use to measure the distance from the capital of the district to the mine. Mineral production is available on an annual

[54] I use the coordinates of the district's capital as provided by the Ministry of Education.

basis from 1975, and by 2013 there were 367 active projects (some mines are involved in several projects).

I use 27 mines related to large-scale projects[55] for the analysis[56]. These were mines active in the baseline years 1993-1994. I use their production in 1993 to construct a weight for the change in mineral prices from 1993 to 2007. I constrain the sample to mines extracting five minerals: gold, silver, copper, lead and zinc. For each mine I, then, construct a measure of the monetary value of production in 1994 for which I use data on international prices reported by the U.S. Geological survey (see Kelly and Matos; 2013). All prices have 1994 as base year. Figure 3.1 shows that most of Peru's GDP is highly influenced by mining production.

Figure 3.1: National and Mining GDP (Thousands of 2007 Soles)



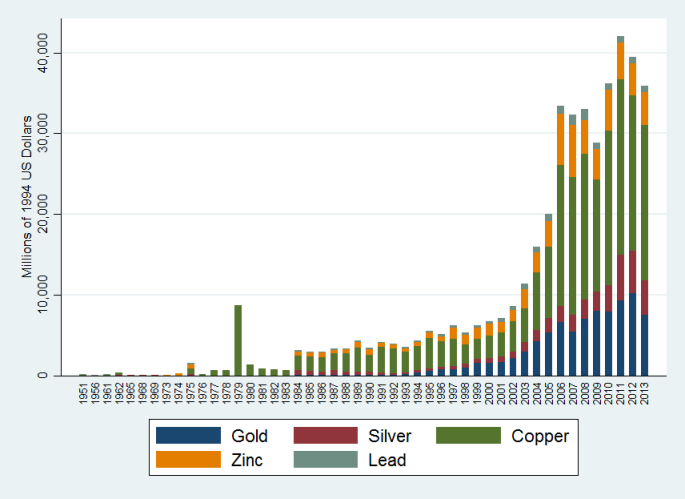
Notes: [1] Data source: Peruvian Central Bank. [2] Red line marks the beginning of the period of analysis.

The annual mining production estimated using data from InterraRMD (2013) from 1950 to 2013 and its composition is presented in figure 3.2. Mineral production boomed from early 2000's. Copper was by large the main mineral extracted from Peruvian mines. Gold production gained importance from middle 1990's.

[55]The data do not contain information on informal mining activity. Informal Mining remains a topic for future analysis. In some regions (which are not part of my empirical analysis) of the country, illegal gold mining boomed in recent years and has brought deforestation and other social consequences. https://www.theguardian.com/environment/andes-to-the-amazon/2016/may/01/gold-mining-in-peru-forests-razed-millions-lost-virgins-auctioned

[56]In between 1993 and 2007 41 new mines started operations, but do not include them because I am interested in the medium or long term effect of the change in mineral prices. The empirical section provides more detail on the selection process.

Figure 3.2: Mining Production (Thousands of 1994 US Dollars)



Notes: [1] Data source for production: InterraRMD (2013). Data source for prices: Kelly and Matos (2013).

To give an idea magnitude of the boom, from 1985 to 1995 mining production for these five minerals grew at an annual average of 3.3%. For the second half of the 1990's the average annual growth rate was 7.3%. During the 2000's the annual growth rate increased to 16%.

Census and household Data

Population censuses are collected by the National Institute of Statistics (INEI in Spanish) on an irregular basis. The last population census dates to 2007, while the closest one dates to 1993[57]. The administrative classification of Peru considered 1,791 districts in 1993. This number increased to 1,831 by 2007, but in order to retain comparability between the two censuses, for the new districts created by 2007 I assign them their administrative code as it was in 1993.

The empirical section selects a group of 1043 districts that are comparable in 1993. Ideally, they would be comparable in levels and in their trends before 1993. In order to confirm their comparability in trends, I collect information from the 1981 census. Unfortunately, the 1981 census is incomplete, and I can only use information from 22 regions out of the 25 that are

[57]Both are *de facto* type of census. INEI collected a *de jure* type of census in 2005, which I do not use to avoid problems of comparability. The distinction between the two types of censuses relates to the location of the individuals at the moment of the interview. A *de jure* type of census enumerates individuals as of where they usually reside. In contrast, a *de facto* type of census enumerates the individual at the place where they were found.

in Peru[58]. This results in a reduction of the number of districts to 931 for which I am able to construct employment and population variables in 1981. In consequence, I use these 931 districts only for pre-trend checks and not for the estimation of the effect of the boom in mineral prices from 1993 to 2007.

For the information on wages I use individual level data. INEI also collects the ENAHO every year since 1997. The survey has changed through time, however it retains fixed many of its important modules from the beginning. For the estimation of the effect of mining activity on wages I use data from the modules on education, employment and income, household characteristics, and individual characteristics. Unfortunately, the survey is statistically representative at the regional level and not at the district level. Not all the districts of Peru are covered every year, however through the sample, the different sample designs of the survey managed to collect information of all districts. In consequence, I rely on estimates at the individual level.[59]

Other data

Through the analysis I also make use of data on altitude, rainfall and land extension. Data on altitude and land extension were available from the Ministry of Energy and the agricultural census[60]. The source for the historical rainfall data is the Climatic Research Unit from the University of East Anglia[61]. I use the version TS3.20 that covers the period from January 1901 to December 2011 and provides precipitation estimates at the 5°×5° grid resolution, that I match to the district borders.

3.3.2 Measuring Migration

The theoretical and empirical analysis is focused on the effects of the boom in the mineral prices over the labor market: employment and wages. However, implicitly there is a migration dimension in the determination of the results.

[58] One of the missing regions falls within the geographic area I use for the empirical resign

[59] Which resembles the strategy adopted by Dell (2010) when assessing the impact of the historical institution of the mining *Mita* over household consumption in Peru.

[60] The agricultural census covers only districts with agricultural land, and therefore excludes a few districts in the coastal region with no agricultural land. This is not a problem since the identification strategy excludes districts from the coast.

[61] Available at http://www.cru.uea.ac.uk/

In this brief section I outline a descriptive exercise aimed to explore with more detail the migration response. Moretti (2010) points out that the literature on local labor markets has not explored the consequences of local shocks into the the type of individual: are local residents or migrants those benefiting from the new context?

The census data typically records two questions related to migration: the district of birth and the district of residence five years ago. The combination of these two categories yields four types of individuals that live in the district. If district i's current labor force is N_i, it can be organized in four migration categories:

$$N_i = M_{i11} + M_{i12} + M_{i21} + M_{i22} \tag{3.1}$$

Where M_{i11} are those individuals who were born in the district and were living there also five years ago, who I term *Locals*. $M_{i12}$are individuals also born in the district but were not living there five years ago, termed as *Returned*. Individuals who were not born in the district but were living there since five years ago, M_{i21}, are classified as *New Locals*. Finally, individuals who were born in a different district and were also living in a different district five years ago, M_{i21}, are *New Comers*.

For each of these categories, I can also explore the response to the natural resource shock by industry and skill groups. Empirically, I explore the effects on migration of the 1043 districts I observe in the analysis. Migration, however, could be to anywhere in the country.

3.3.3 Empirical Framework

The unit of observation for the mining activity is the district. Ideally I would like to observe differences in the employment, population and wage variables at the district level when a district with mining activity benefits from the surge in the international prices of the minerals the firms extract. However, I can only retrieve employment and population (and migration) at the district level, while for the estimation of the effect on wages I end up using individual level data. The identification strategy selects a set of districts that are comparable at the baseline, 1993. Then it compares their response to the boom in commodity prices from 1993 to 2007.

In 1993 the 27 mines extracted the following minerals: gold, silver, copper, zinc and lead. One mine extracted the 5 minerals, 5 mines extracted 4 minerals, 11 mines extracted three minerals, 4 mines extracted 2 minerals, while 6 mines extracted only one mineral. This feature is going to be exploited in the identification strategy (i.e. each treated district was exposed not only to different number of mines but also different production of minerals by each mine).

By 2007 the number of mines had increased to 68. Although new mines offer an interesting analysis, many of them started operation between 2004 and 2006, which puts them closer in time to the end-line of the time frame, and would be more suitable for a short-term analysis. In this chapter I am interested in the medium or long-term differences. Therefore, in order to rule out any interference coming from these 41 new mines, I excluded those districts that were not under the influence of a mine in 1993 but became influenced by 2007.

In 1993, I consider that a district was under the influence of any of the 27 operating mines if the distance between its capital and any of the mines was less than 100km. This threshold is assumed in the literature and I opt for it[62]. For the estimation of the distance I use the length of the shortest curve between the coordinates of the capital of the district, and the coordinates of the closest mine[63]. I refer to them as mining districts. Then I locate districts who's capital was more than 100km apart from any of the 27 mines, but less than 200km. These are neighboring districts that could be understood as *control* group[64]. The idea is that by 1993 there were no meaningful differences in observables between districts with a mine at a 100km distance and the closest neighboring districts in the 200km radius of the mine. The choice of 100km can also be defended by the fact that the average district in Peru has an area of 71,866 Km2 while the median district has an area of 20,852 Km2. This means that the 100 km threshold is slightly below the square root of the median area.

There were a few additional steps before reaching the final number of districts in both groups. As mentioned previously, I also excluded any district with an altitude below 1,800 meters above the sea level. The result of this is the exclusion of coast cities, and cities located

[62]For instance, in a similar setting, Aragón and Rud (2013) define 100 km as the influence distance. Fafchamps et al. (2015) assume no effect of mining activity beyond 100 km.

[63]I implement this by using the STATA command `geodist`.

[64]As mentioned above, I excluded districts that were part of the control group in 1993 but became mining districts in the period following up to 2007.

in the rain-forest. There are two reasons for this. First, Peru's geography is very heterogeneous and exerts an important influence in the markets[65][66]. Second, there are no large-scale mines operating below 2,000 meters above the sea level. Therefore, the 1,800 threshold allows the inclusion of downstream cities that are still under the influence of the mining activity but somehow isolated from the coast or rain-forest dynamics.

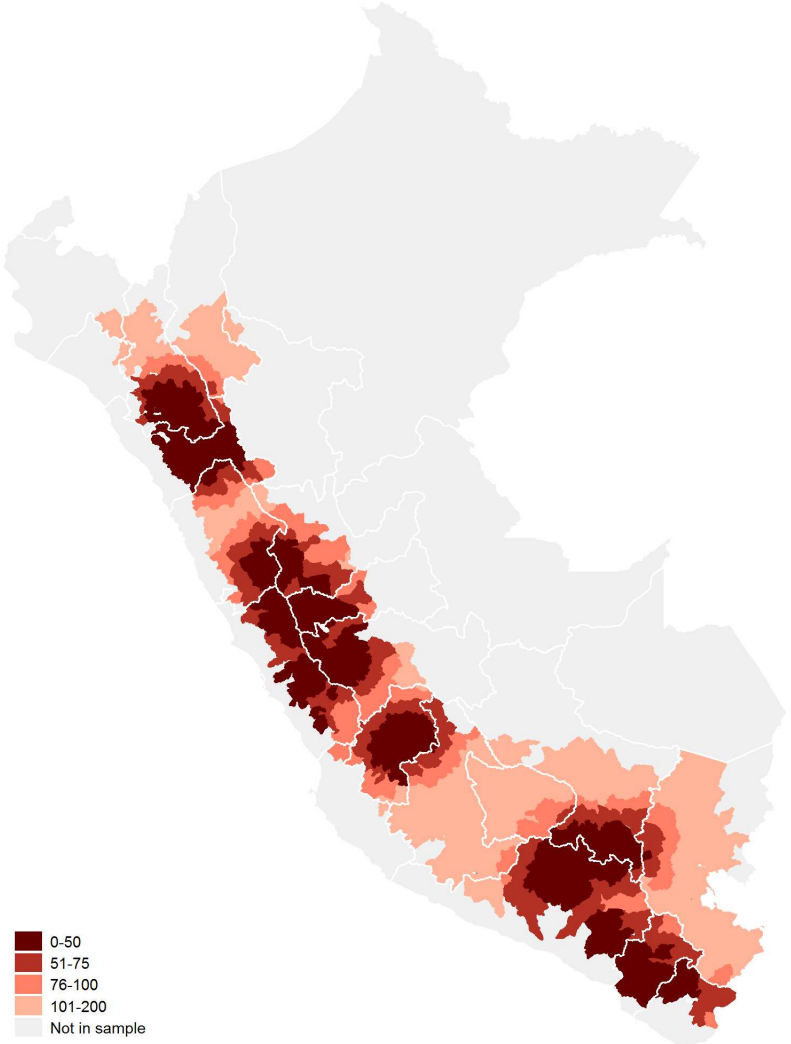
After all this criteria, the total number of districts included in the analysis was 1043. 823 districts were mining districts, within the 100km distance range from a mine, while 220 control districts were within the 101km-200km distance interval and never saw a new mine opening in subsequent years[67]. Figure 3.3 illustrates the region under analysis. No district from the coast or the jungle is included. The design allowed the inclusion of districts from the north, central and south highlands. I preset the color ranges based on distance to provide an idea of the district's proximity to mines, but I consider districts as treated whenever it had a mine within the 100km range or below.

[65]The *Andes* mountains divide the country into two parts, the coast on the West and the rain-forest on the East, leaving the *Andes* as a middle region. This partition influences mainly transport costs, which ultimately explains the differences in productivity as documented by Sotelo (2015).

[66]Former Peruvian president Manuel Pardo, during the early days of the Peruvian Republic in 1862, coined an illustration that has survived to this day: the freight cost from Jauja (a small city in the central highlands) to the capital, Lima was four times higher than the maritime freight from the capital harbor, Callao, to Liverpool. See Pardo (1996)

[67]As noted before, 41 of the original *control* districts were influenced by the opening of new mines. I removed these districts.

Figure 3.3: Mining and Non-Mining Districts



Notes: [1] Red coluring intensifies with proximity to any of the 27 mines operating in 1993. Light red represents the districts not affected by mining activity (within 101km and 200km). [2] Grey area represents districts that are not considered in the analysis.

The proximity to the mine is not the only difference between districts. The empirical strategy actually exploits the fact that mines had different portfolios of minerals in order to construct a measure of the price boom experienced by each mining districts. I do this in two stages. First, I estimate the weighted average price change (from 1993 to 2007) for every mine:

$$\Delta P_g = \sum_m \kappa_m \Delta P_m \tag{3.2}$$

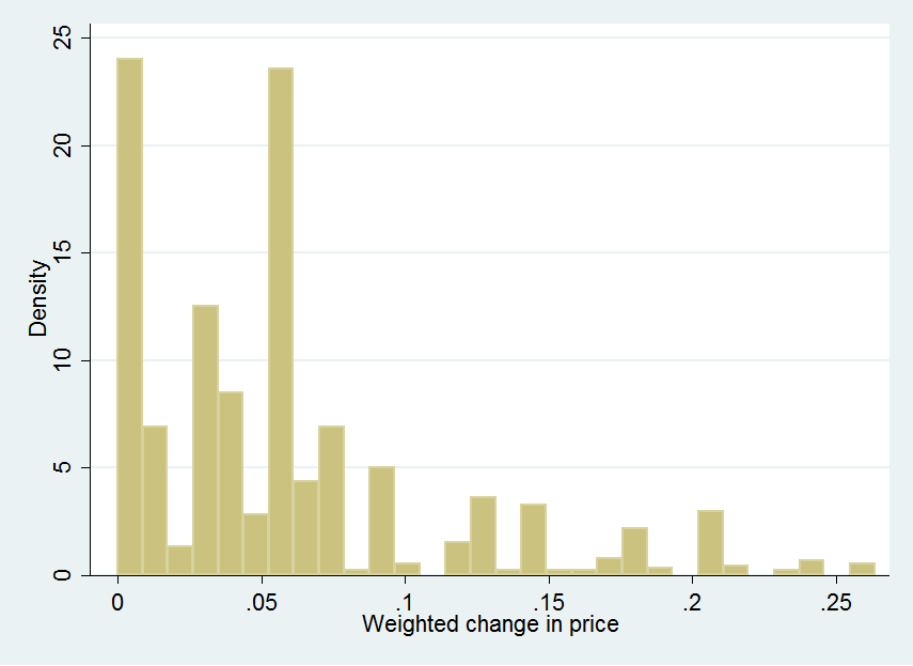
Where m={gold, silver, copper, zinc, lead}. κ_m is the mineral weight, estimated as the mineral production in 1994 valued at 1994 international prices. ΔP_m is the international price change (increase in percentage points) in the mineral from 1993 to 2007. g index each of the 27 mines. This measure exploits variability in the mineral portfolio of each mine.

In the second stage, I match every district in 1993 to all the mines that fell within the 100km distance threshold. Some districts were under the influence of many mines, therefore, the final measure of price boom adds one additional source of variation: the number of mines influencing the district. In concrete, the district measure that I construct is:

$$\Delta P_i = \sum_g \Delta P_{ig} \tag{3.3}$$

Where each district i was under the influence of G_i mines, each with a mineral portfolio that yielded a price change ΔP_g. The measure ΔP_i is simply the sum of price changes over all mines that influenced the district. With this measure, districts surrounded by several mines will experience heavier influence of mining than districts close to one mine. Figure 3.4 plots the distribution of ΔP_i.

Figure 3.4: Distribution of ΔP_i



The empirical equation, therefore, for the effect on labor N in district i is:

$$\Delta(N_i) = \alpha + \beta \Delta P_i + \theta_k \Gamma_{ki} + \eta_p + \varepsilon_{it} \tag{3.4}$$

Where $\Delta(N_i) = \ln(N_{i2007}) - \ln(N_{i1993})$, which is the percentage change in employment in district i. I subsequently evaluate the effect by industry and skill group[68]. However, I also evaluate a different version: $\Delta(N_i) = (N_{i2007}/Pop_{i2007}) - (N_{i1993}/Pop_{i1993})$, which is the change in the employment rate, where $Pop_{i,1993}$ is always the district adult population (16 years old or more). In consequence, empirically, the first definition evaluates differences in the growth rate of the number of workers, while the second evaluates the differences in the change in the employment rate. Again, I extend this estimation to the industry and skill groups.

Having the dependent variable as change in rates as a complement to the change in (log) number of workers is an attempt to understand the re-shuffling of employment in the context of a geographical model with constant population. Then, to directly understand how much of the hypothetical change is explained by new high skilled or low skilled workers, the employment rate is separated by the share explained by those two categories. For instance, employment rate in the 1043 districts studied in 1993, was 51.9%. Out of this number, 4.4% was skilled employment, while 47.5% was low skilled employment. The average change in log number of employed people between 2007 and 1993 was -0.035 log points. The high skilled employment changed on average by 0.037 log points, while low skilled employment changed on average by -0.074 log points. In consequence, in this context, finding a positive effect on overall, high skilled and low skilled employment would be interpreted as overall and high skilled employment growing faster, while low skilled employment decreasing slower for those districts under the influence of large-scale mining activity.

The coefficient of interest is β which captures the effect of the boom in mineral prices on the city, as described above. θ_k is a vector with district controls: altitude, the historical coefficient of variation of rainfall, distance to Lima, land extension of the district and the initial level of population (in logs). η_p are province fixed effects. ε_{it} is the empirical disturbance.

[68] I define skilled worker if he or she has at least technical education. This information is available in the census and survey data. With this definition, 18% of the individuals in the survey data are high skilled. The census data indicates that 14% of the population was high skilled in 1993, while this percentage increased to 20% in 2007.

The measure of wages corresponds to all individual earnings from main and secondary occupation expressed on a monthly basis at 2007 prices. All individuals with a dependent job or self-employed were considered. The analysis excludes individuals who report working as unpaid workers. From ENAHO I can also include several individual characteristics as controls. To measure the effect of the boom in mineral prices on wages I try three specifications. The first is simply an application of equation (3.4) at the individual level, which yields the following estimating equation:

$$\ln w_{cit} = \alpha + \beta \Delta P_i + \Omega_a \Pi_{act} + \theta_k \Gamma_{ki} + \eta_p + t_t + \varepsilon_{cit} \tag{3.5}$$

Where $\ln w_{cit}$ is the natural log of monthly wage of individual c living and working in district i at time t. The right hand side of the equation is similar to the employment equation, with the exception that I also control for a set of a individual characteristics Π_{act}: gender (male), age and its square, the number of years of schooling, household size, number of income earners in the household, dummies for water and electricity in the household, fixed effects for industry (two digits ISIC), job type (owner, self-employed, white collar, blue collar), year, t_t, and province, η_p. The variable ΔP_i is the same across time for each district and measures, as usual, the effect of the price change from 1993 to 2007.

The previous equation, however, does not exploit annual variation in the cross section. It does not control for district fixed effects either. To address this, I re-define the mineral price variable as a price index (base 1994) that uses as weights the mineral production of 1994, and is evaluated lagged one period to allow for adjustment: PI_{t-1}[69]. The empirical equation, in this case, take the following form:

$$\ln w_{cit} = \alpha + \beta PI_{it-1} + \Omega_a \Pi_{act} + \eta_i + t_t + \varepsilon_{cit} \tag{3.6}$$

The specification keeps the individual controls, but now geographical fixed effects are at the district level: η_i. Therefore, any change on wages steams from the time variation in the mineral price index.

[69]That is, the weighted mineral price average for every year.

3.4 Results

3.4.1 Employment, population and unemployment

Table 3.1 presents the baseline results for equation 3.4. Panel a. shows the results for the dependent variables as change in log levels, while panel b. evaluates the model with the dependent variable as change in the employment rate. There are statistically significant results for the variables measured as rates, in panel b. The columns on employment (4, 5 and 6) indicate that the increase in mineral prices resulted in higher change in employment rate for the districts under the influence of the mining activity. In particular, in column 4, a one standard deviation of the change of the price index (0.31) is related to a 4% (0.31x0.13) faster change in employment rate for those districts close to any mining activity. This means that, for instance, if both groups of districts experienced an increase in the employment rate, the increase in the mining districts was 4% higher for every standard deviation increase in the portfolio of mineral prices. This effect is large compared to the average change in the employment rate experienced in the sample of districts for the 1993-2007 period. On average, the change in total employment rate was -3%. The result found here suggests that the boom in the mining activity offset the decline in the average employment rate experienced in the sample districts and instead, employment grew by 1%.

Table 3.1: Change in Employment: 2007-1993

	Population			Employment			Unemployment		
	Ln. N.	High Skilled	Low Skilled	Rate	High Skilled	Low Skilled	Rate	High Skilled	Low Skilled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
a. Changes in Log Numbers									
ΔP	-0.09	0.10	-0.15	0.12	0.31	0.05	-0.44***	-0.53	-0.47***
	(0.14)	(0.25)	(0.13)	(0.17)	(0.21)	(0.16)	(0.12)	(0.36)	(0.11)
R^2	0.41	0.37	0.47	0.31	0.36	0.30	0.42	0.34	0.45
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124
b. Changes in Shares									
ΔP		0.05**	-0.05**	0.13***	0.05***	0.07***	-0.13***	-0.00	-0.13***
		(0.02)	(0.02)	(0.03)	(0.01)	(0.03)	(0.03)	(0.01)	(0.04)
R^2		0.52	0.52	0.33	0.52	0.25	0.33	0.37	0.36
Observations		1043	1043	1043	1043	1043	1043	1043	1043
Clusters		124	124	124	124	124	124	124	124

Notes: [1] Data source: population censuses of 1993 and 2007. [2] All regressions include as controls: altitude, historical coefficient of variation of rainfall, distance to Lima (not included in the sample), the log of district population in 1993 and province fixed effects. [3] Panel a. measures the dependent variales in log changes, whereas panel b uses the change in rate. For the estimation of rates the denominator is always the district adult population (16 years old or more). [4] Errors clustered at province level, and coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

Columns 5 and 6 account for employment change by skill type. By doing this, the coefficient for total employment can be broken into the two skill categories. Column 5 indicates that out of the 0.13 percentage points effect, 0.05 is explained by increase in high skilled employment rate, while the remaining 0.07, by low skilled employment (column 6). The three results are statistically significant at 1%.

Regarding population, columns 2 and 3 in panel b, however, indicate that there was a reshuffling in the composition of the population by skill. The estimated coefficient for the change in the share of high skilled population is 5%, which suggests that for every additional standard deviation of the mineral prices, the share of high skilled individuals increased by 1.6%. This result is not small. The average percentage of high skilled people in both periods was 7.3%, while the average change in the share of high skilled individuals was 3.7%. The estimated effect indicates, therefore, that for every standard deviation increase in the mineral prices, mining districts experienced an additional increase that was 43% (0.016 out of 0.037) higher than the average increase in in the share of high skilled population.

Panel a offers no statistically significant result either for population or employment, however the sign of the estimated coefficients are in line with the results in panel b. Which suggests that more than in absolute terms, mining had an effect on the structure of the labor market.

Columns 7, 8 and 9 evaluate any impact on unemployment. Total unemployment decreased in districts close to the mining activity measured both as the change in log number of unemployed individuals (panel a) or as a share (panel b). The reduction of unemployment is particularly evident for low skilled workers. From panel a, mining districts experienced a faster decline in growth rate of the number of unemployed individuals by a magnitude of 14% (0.47x0.31) by every standard deviation in the mineral price index. As a share (panel b), mining districts saw low skilled unemployment rate decline by 4% (0.13x0.31) faster than in non-mining districts.

Results hitherto suggest that the mining boom experienced from 1993 to 2007 had effects on the employment rate of the local economy. This increase in employment reached both skilled and unskilled workers, with a certain emphasis on skilled workers. The proportion of skilled individuals also increased in the local economy close to the mine, while unemploy-

ment for the low skilled group dropped. Employment results in columns 5 and 6 of panel b are in line with the theoretical predictions under the assumption that skilled workers are more mobile than unskilled workers.

Table 3.2 evaluates the heterogeneous effect on employment by industries. Panel (a) evaluates the effect on the growth rate of employment (difference in logs), and indicates that employment in the mining sector grew faster. The effect on this sector is quite large. The average increase in mining employment in the districts included in the sample between 1994 and 2007 period was 84%. The coefficient for the effect on overall mining employment reported in column 1, 0.89, indicates that by every standard deviation increase in the price of minerals (0.31), the mining districts experienced an additional increase in employment of 28% (0.89x0.31). This effect was stronger for skilled employment. The average increase in skilled employment between 1993 and 2007 was 61%. The estimated coefficient for the effect on skilled employment was 1.00, which interpreted in terms of standard deviations reports an extra increase in skilled employment of 31% (1.00x0.31) in the mining districts. The rest of the coefficients in panel (a) are not statistically significant.

Table 3.2: Change in Employment: 2007-1993, By Industry and Skill

	Mining			Tradable			Non-Tradable			Agriculture		
	N	H	L	N	H	L	N	H	L	N	H	L
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
a. Changes in Log Numbers												
ΔP	0.89*	1.00***	0.70	0.08	-0.12	0.04	-0.01	0.20	-0.17	0.08	-0.02	0.06
	(0.48)	(0.37)	(0.45)	(0.20)	(0.24)	(0.19)	(0.22)	(0.23)	(0.22)	(0.17)	(0.27)	(0.16)
R^2	0.29	0.31	0.29	0.26	0.18	0.26	0.30	0.29	0.29	0.27	0.27	0.27
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
b. Changes in Shares												
ΔP	0.05**	0.03***	0.01	-0.00	-0.00	-0.00	0.01	0.01	0.00	0.06**	0.01	0.05**
	(0.02)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)
R^2	0.18	0.23	0.15	0.21	0.27	0.22	0.33	0.49	0.22	0.29	0.48	0.28
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124

Notes: [1] Data source: population censuses of 1993 and 2007. [2] All regressions include as controls: altitude, historical coefficient of variation of rainfall, distance to Lima (not included in the sample), the log of district population in 1993 and province fixed effects. [3] Panel a. measures the dependent variales in log changes, whereas panel b uses the change in rate. For the estimation of rates the denominator is always the district adult population (16 years old or more). N holds for total, H for high skilled workers and L for low skilled workers. [4] Errors clustered at province level, and coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

Panel (b) reports the changes in the employment rate. The estimated coefficients are consistent with the idea of the mining boom increasing the employment rate across sectors. To

understand these coefficients it is worth recalling the way they are defined. Each employment category (industry, or industry-skill type) is represented as a share of total employment. In consequence, the total effect on employment reported in table 3.1 (0.13; panel (b) column 4) is divided between the four sectors considered in the analysis. In that regard, the results show that 84.6% of that effect was driven by increase in the employment share in mining (0.05, column 1) and agriculture (0.06, column 10). More interestingly, out of the 0.05 percentage points effect in the mining employment, 0.03 are explained by and increase in the high skilled employment. Whereas in the agriculture sector, out of the 0.06 percentage points increase, 0.05 is explained by low skilled employment. This result, therefore, confirms the heterogeneous nature of the effect of the boom in mineral prices at both levels: sectors and skills.

The no effect on tradable and non-tradable industries is in itself an interesting result. The idea that the tradable sector may be particularly hurt by a boom in an industry that increases the costs of factors but not the price of the traded goods finds no validation in the current set of results.

In order to confirm the validity of the results, tables 3.3 and 3.4 conduct a falsification test. In table 3.3 I evaluate whether the population composition and employment measured as rates before the boom in mineral prices were affected by future boom. Panel (a) of table 3.3 evaluates the population and employment variables as in 1993 (all variables measured as rates except for column 1 which reports the log number of population and column 4 which reports results for the change in log employment). There is no result statistically significant in panel (a): the spatial difference between mining and non-mining districts was not statistically significant. Panel (b) proceeds with the same evaluation, but using the change from 1981 to 1993 in the dependent variables. For this evaluation, however as explained above, I can only use 22 regions, which reduces the number of districts to 931. Results are reassuring, and the future mining boom is not associated with previous trends in population or employment, except for the employment rate which is negatively associated with the future mining boom. However, if only non-mining employment is considered, the placebo test provides the expected result. At this point results at least are reassuring for all variables but those related directly to the mining. It is also possible that the missing information from the 112

missing districts would validate the placebo test. I will leave this to future research.

Table 3.3: Balance: Demographics in 1993

	Population			Employment				Non-Mining Employment		
	Ln. N.	High Skilled	Low Skilled	Ln. N.	Rate	High Skilled	Low Skilled	Rate	High Skilled	Low Skilled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
a. Dependent variables in shares in 1993										
ΔP	-0.30	0.01	-0.01	-0.41	-0.07*	-0.00	-0.07	-0.08*	-0.01	-0.07
	(0.68)	(0.04)	(0.04)	(0.65)	(0.04)	(0.02)	(0.06)	(0.04)	(0.02)	(0.06)
R^2	0.54	0.36	0.36	0.53	0.31	0.31	0.31	0.32	0.29	0.32
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124
b. Dependent variables change in shares 1993-1981										
ΔP	0.02	0.01	-0.01	-0.11	-0.09**	-0.00	-0.08**	-0.04	-0.00	-0.04
	(0.21)	(0.01)	(0.01)	(0.25)	(0.04)	(0.01)	(0.04)	(0.04)	(0.01)	(0.04)
R^2	0.38	0.46	0.46	0.32	0.20	0.45	0.17	0.20	0.46	0.17
Observations	931	931	931	931	931	931	931	931	931	931
Clusters	112	112	112	112	112	112	112	112	112	112

Notes: [1] Data source: (i) population census of 1993 for panel a. (ii) population censuses of 1993 and 1981 for panel b. [2] 1981 census has fewer districts because, according to the public information provided by INEI, it was not possible to recover data from three regions (Apurimac, Loreto and San Martin). Population of these three regions represent 6 % of total population in 1981. Moreover Loreto region, and partially San Martin, fall outside the sample of districts used in this analysis. see INEI (2015). [3] All regressions include as controls: altitude, historical coefficient of variation of rainfall, distance to Lima (not included in the sample) and province fixed effects. [4] Column 1 is the natural log of district population (above 16 years old). Columns 2 and 3 are the share of high skilled and low skilled individuals. Columns 4 to 7 refer to employment. Column 4 is the natural log of the number of workers in the district, column 5 is the employment rate, columns 6 and 7 are the number of high and low skilled workers ove the population. Columns 8-9 replicate exclude mining workers. [5] Errors clustered at province level, and coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

Table 3.4 directly explores whether the boom in mineral prices between 1993-2007 explained the change in the employment rate by industry and skill group from 1981 to 1993. Again, it is not the case, although the estimations include only 931 districts.

Table 3.4: Balance: Pre-Trends in log level employment: 1993-1981

	Tradable			Non-Tradable			Agriculture		
	N	H	L	N	H	L	N	H	L
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ΔP	0.00	0.00	0.00	-0.02	-0.00	-0.01	-0.03	0.00	-0.03
	(0.01)	(0.00)	(0.00)	(0.02)	(0.00)	(0.02)	(0.04)	(0.00)	(0.04)
R^2	0.22	0.34	0.21	0.20	0.44	0.20	0.16	0.27	0.16
Observations	931	931	931	931	931	931	931	931	931
Clusters	112	112	112	112	112	112	112	112	112

Notes: [1] Data source: population censuses of 1993 and 1981. [2] 1981 census has fewer districts because, according to the public information provided by INEI, it was not possible to recover data from three regions (Apurimac, Loreto and San Martin). Population of these three regions represent 6 % of total population in 1981. Moreover Loreto region, and partially San Martin, fall outside the sample of districts used in this analysis. see INEI (2015). [3] All regressions include as controls: altitude, historical coefficient of variation of rainfall, distance to Lima (not included in the sample) and province fixed effects. [4] All dependent variables are the change in the employment rate of the industry, by skill type. In all cases, the denominator is the total population whle the numerator is the number of workers in the corresponding industry and industry-skill pair. N holds for total, H for high skilled workers and L for low skilled workers. [5] Errors clustered at province level, and coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

I tested the same variables using the the log number of individuals instead, and results were even more reassuring since I found no effect at all in the placebo test. For space considerations I did not report the tables.

3.4.2 Effect on wages

Results so far indicate that high skilled workers found more jobs in districts under the influence of the mineral price boom. This also meant that the share of individuals with high skill increased while the proportion of low skilled individuals decreased. The theoretical framework predicts that under this circumstance the wage for low skilled workers should experience a higher increase than the wage for high skilled workers.

Table 3.5 presents the results of estimating equations (3.5) in panel a., and equation (3.6) in panel b. The interpretation of these equations are slightly different than the employment estimations. In this case, the boom of mineral prices explains wage differentials in the log levels of the two groups of districts, rather than the growth differences.

I use 1997-2007 for the estimation. The government started producing ENAHO since that year, therefore I cannot collect individual wage information for years before 1997.

Column 1 estimates the effect on total monthly wage, and finds a positive effect in all specifications. The coefficient shown panel (a), 0.17, suggests that every standard deviation increase in the mineral prices was associated with an average wage 5.2% (0.17x0.31) higher in

Table 3.5: Wages by Skill: 1997-2007

	Both	High Skilled			Low Skilled		
		>= Uni.	>= Tec.	>= Sec.	< Uni.	< Tec.	< Sec.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
a. ΔP							
ΔP	0.17***	0.03	0.16*	0.13***	0.19***	0.19***	0.22***
	(0.04)	(0.09)	(0.09)	(0.05)	(0.04)	(0.05)	(0.06)
R^2	0.51	0.48	0.46	0.46	0.47	0.44	0.39
Observations	73272	7467	14511	27750	65791	58747	45508
Clusters	702	381	508	648	702	702	702
b. PIt							
PI_{it-1}	0.38*	-0.01	-0.15	-0.22	0.42**	0.53**	0.72***
	(0.20)	(0.47)	(0.31)	(0.22)	(0.21)	(0.22)	(0.24)
R^2	0.53	0.50	0.48	0.48	0.49	0.46	0.42
Observations	73272	7467	14511	27750	65791	58747	45508
Clusters	702	381	508	648	702	702	702
c. Production							
$Production_{it-1}$	0.11***	0.13**	0.07**	0.13***	0.10***	0.11***	0.08**
	(0.03)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
R^2	0.53	0.50	0.48	0.48	0.49	0.46	0.42
Observations	73272	7467	14511	27750	65791	58747	45508
Clusters	702	381	508	648	702	702	702

Notes: [1] Data source: ENAHO 1997-2007. [2] Dependent variable is the log of monthly wage of main and secondary ocupation at 2007 prices. Column 1 considers all individuals. Columns 2, 3 and 4 use data from high skilled workers. High skilled are those individuals with university education (column 2), or at leat technical education (column 3), or at least secondary education (column 4). Columns 5, 6 and 7 refer to low skilled wages as the alternative to columns 1, 2 and 4, respectively. [3] Panel a. replicates the empirical equation used in the district level regression. Panel b. estimates the effect on wages using a yearly commodity price index. Panel c. uses annual mining production. Individual controls included in panels b. and c.: gender (male), age and its square, the number of years of schooling, household size, number of income earners in the household, dummies for water and electricity in the household, fixed effects for industry (two digits ISIC), job type (owner, self-employed, white collar, blue collar), year and district. [4] Errors clustered at province level in panel a. Errors clustered at district level in panels b. and c. Coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

the mining districts. Panel (b) estimates the equation controlling for district and year fixed effects, and uses an annual price index, instead. The effect of a one standard deviation increase in this case is higher: mining districts reported a monthly wage 11.8% higher (0.38x0.31[70]). Using annual production, as in panel (c), the estimated coefficient is 0.11, which represents a wage 8.6% (0.11x0.78[71]) higher in mining districts. The estimates of panel (c), however, may be less defensible in terms of endogeneity. Annual mineral production depends on the amount of labor and capital available every year and attracted to the local economy. In that regard, results from panels (a) and (b), are more reliable: they use 1994 production as weight for the index. Panel (a) uses the 1994-2007 change in prices, while in panel (b) the estimation uses an index constructed with 1993-1994 productivity weights.

Columns 3 and 6 show the estimated effect of the mineral price boom on the monthly

[70]Mean and standard deviation of the annual index for mineral prices: 0.29 and 0.31, respectively
[71]Mean and standard deviation of annual production: 0.79 and 0.78, respectively

wage by the skill type. As predicted by the theory, the wage for low skilled workers should experience a higher increase than the wage for high skilled workers. This is reflected in column 6, panel (a), using the baseline definition of high skill: technical studies[72]. The estimated coefficient is 0.19, which is 3 percentage points higher than the estimated effect for high skilled wage (0.16 in column 3). However these two coefficients are not statistically different.

How does the skill threshold affects the estimates? A more strict definition of high skill, university degree[73], makes the result clearer, and this is reflected in columns 2 for high skilled wage and 5 for low skilled wage. With this definition, the effect on low skilled wages is the only one statistically different from zero, which ultimately implies that this type of workers are the only wage gainers from the productivity shock in the natural resource sector.

With a less strict threshold for skill (secondary education), the estimated effects are again not statistically different (although the higher point estimate for low skilled wage): the estimated effect for low skilled wage is 0.22, with an standard error of 0.06, while the estimated effect for high skilled wage is 0.13 with an standard error of 0.05.

The estimation procedure that consider the annual price index, reported in panel (b), does indicate that low skilled wages reacted positively in mining districts, while high skilled wages did not. More interestingly, the point estimate for the effect on low skilled wage increases for lower thresholds in the skill definition. In concrete, for the lowest threshold that considers an individual as low skilled if his or her qualification is smaller than secondary (that is, individuals with primary education and no education) yields an estimated effect of 0.72. This effect suggests that low skilled workers in mining districts had a wage 22% (0.72x0.31) higher than the wage of low skilled workers in non-mining districts.

Estimates from panel (c) show estimated coefficients for both skill groups that are not statistically different from each other.

Results on wages so far do not clearly reflect a difference response by skill type. It is only on panel (b) that low skilled wages reacted positively to the boom in mineral prices. In table 3.6 I shed more light on the heterogeneous wage response by industry. The estimation

[72]I assumed technical studies as the threshold for skilled labor. But I evaluate the sensitivity of this threshold later in this section.

[73]With this definition, 9% of the individuals in the survey sample are high skilled.

procedure is the same, but in this case I constrain the sample of individuals depending on the industry the belong into. Panel (a) suggests that there is only a detectable effect on the monthly wage of agricultural workers. Surprisingly, there is no effect on the wage of mining workers, but this may be a consequence of the reduced number of individuals working in the mining sector in the 10 years period from 1997 to 2007 included in the sample: 1,686. The estimated coefficient reported in column 4, 0.25, suggests that the average monthly wage of agricultural workers in mining districts was 7.8% higher that the average monthly wage of agricultural workers in non-mining districts by every standard deviation increase in the price index for minerals. The effect on agriculture found in panel (a) can be explained by the large share of low skilled workers in that sector. Only 3% of the individuals in the agricultural sector have at least technical education, whereas 73% have no education or primary education.

Table 3.6: Wages by Industry: 1997-2007

	Mining	Tradable	Non-Tradable	Agriculture
	(1)	(2)	(3)	(4)
a. ΔP				
ΔP	-0.27	-0.06	0.03	0.25***
	(0.23)	(0.13)	(0.06)	(0.08)
R^2	0.67	0.61	0.44	0.29
Observations	1686	6402	34458	30726
Clusters	187	516	672	696
b. PIt				
PI_{it-1}	0.25	0.70	0.00	0.47
	(0.64)	(0.54)	(0.18)	(0.32)
R^2	0.73	0.65	0.46	0.33
Observations	1686	6402	34458	30726
Clusters	187	516	672	696

Notes: [1] Data source: ENAHO 1997-2007. [2] Dependent variable is the log of monthly wage of main and secondary ocupation at 2007 prices, by indistry type. [3] Panel a. replicates the empirical equation used in the district level regression. Panel b. estimates the effect on wages using a yearly commodity price index. Individual controls included in panel b. : gender (male), age and its square, the number of years of schooling, household size, number of income earners in the household, dummies for water and electricity in the household, fixed effects for industry (two digits ISIC), job type (owner, self-employed, white collar, blue collar), year and district. [4] Errors clustered at province level in panel a. Errors clustered at district level in panel b. Coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

Panel (b) finds a positive effect for mining, tradable and agriculture, but none of the coefficients is statistically significant. Panel (c), which shows the results for an estimation that uses annual production, suggests that the wages in the mining, tradable and agricultural sectors were higher in mining districts. However, this estimation is less reliable due to potential problems of endogeneity.

3.4.3 Effect on migration

So far the results focused on employment and wages. The estimated effects are in line with the theoretical predictions of higher employment, with an emphasis on high skilled workers, and therefore a higher increase in low skilled wages. In this section I address the explanation of the results from the migration perspective. However, I adopt a descriptive point of view here.

As mentioned in equation 3.1, the current population, or labor force can be categorized in four groups, indexed by m: *locals*, *new locals*, *individuals who returned*, *and new comers*. To investigate the effect of the mining boom in any of these categories, I estimate the following equation:

$$\Delta\left(M_{mi}\right) = \alpha + \beta\Delta P_i + \theta_k \Gamma_{ki} + \eta_p + \varepsilon_{it} \tag{3.7}$$

This equation is the same as (3.4) but uses any of the migration categories as dependent variable. Again, as in the case of population and employment, I measure the effect of the mining boom over two different types of dependent variables. First, $\Delta\left(M_{mi}\right) = \ln M_{mi,2007} - \ln M_{mi,1993}$ represents the log change in the number of migrants in category m in district i[74]. Second, I also measure the effect of the mining boom over $\Delta\left(M_{mi}\right) = M_{mi,2007}/Pop_{i,2007} - M_{mi,1993}/Pop_{i,1993}$, which in this case is defined as a share of the working population[75].

Table 3.7 shows the results. Panel (a) presents the results using the change in the log number of individuals for population (panel a.1) and labor force (panel a.2), while panel (b) groups the results for the migration rates measured as population share (panel b.1) of employment share (panel b.2)[76].

[74] Later also considered by industry and skill group.

[75] Individuals 16 years old or more

[76] For example, when measuring the proportion of locals in panel b.1 I use the total number of locals (employed or not) over the working population; while in panel b.1 I use only the locals who are employed in the numerator, leaving the denominator the same.

Table 3.7: Change in Migration Composition: 2007-1993

	N22 Locals			N21 Returned			N12 New Locals			N11 New Comers		
	N	H	L	N	H	L	N	H	L	N	H	L
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
a. Changes in Log Numbers												
a.1 Population												
ΔP	-0.11	-0.04	-0.17	0.02	-0.03	-0.00	-0.00	0.22	0.02	0.30	0.45	0.27
	(0.12)	(0.23)	(0.11)	(0.17)	(0.17)	(0.18)	(0.14)	(0.21)	(0.13)	(0.29)	(0.33)	(0.27)
R^2	0.77	0.44	0.78	0.41	0.25	0.39	0.44	0.29	0.44	0.34	0.28	0.33
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
a.2 Employment												
ΔP	0.09	0.25	0.02	0.20	0.01	0.27	0.09	0.39*	0.08	0.33	0.37	0.39
	(0.14)	(0.19)	(0.13)	(0.14)	(0.14)	(0.17)	(0.18)	(0.21)	(0.16)	(0.27)	(0.29)	(0.26)
R^2	0.62	0.37	0.61	0.37	0.24	0.35	0.36	0.27	0.34	0.32	0.31	0.29
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
b. Changes in Shares												
b.1 Population												
ΔP	-0.00	0.03	-0.04	-0.01	-0.00	-0.01	-0.01	0.01	-0.02	0.02	0.02*	0.01
	(0.02)	(0.02)	(0.03)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.47	0.50	0.50	0.32	0.20	0.34	0.15	0.37	0.15	0.72	0.26	0.79
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
b.2 Employment												
ΔP	0.09***	0.03***	0.06**	0.00	-0.00	0.00	-0.00	0.01**	-0.01	0.02*	0.01**	0.01
	(0.03)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.34	0.50	0.30	0.31	0.17	0.34	0.20	0.45	0.17	0.49	0.31	0.56
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124

Notes: [1] Data source: population censuses of 1993 and 2007. [2] All regressions include as controls: altitude, historical coefficient of variation of rainfall, distance to Lima (not included in the sample), the log of district population in 1993 and province fixed effects. [3] The dependent variables are the current population composition (share) of the district according to migration categories. *Locals* are those individuals born in the district that were living there five years ago. *Returned* are individuals born in the district who were not living there five years ago. *New Locals* are individuals not born in the district but were living there five years ago. *New Comers* are individuals who were not neither born or residents of the districts five years ago. For each category N holds for total, H for high skilled workers and L for low skilled workers. [4] Errors clustered at province level, and coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

Focusing on panel a.2, which presents the results for the the change in log employment, there is only a 10% significant increase for the log number of high skilled *new locals* for mining districts (column 8). In population terms, the number of low skilled locals (column 3, panel a.1) seems to decline, but this result is not statistically significant.

Panel b evaluates the changes in each migration category as share of population and employment. There is no discernible difference across migration categories to explain the re-shuffle of population (panel b.1). However, in terms of employment (panel b.2), the effect over the proportion of locals is 0.09 percentage points (column 1). This coefficient bear an important result. In table 3.1 the estimated effect of employment was 0.13, therefore, the coefficient 0.09 estimated here indicates that 0.09 out of the 0.13 points increase in employment is explained by more employment of locals, this represents 69%. Column 10 in panel b.2 reports a coefficient of 0.02 which bears similar interpretation: 0.02 percentage points out of the 0.13 percentage points effect found in the baseline results are explained by new comers, which represents 15% of the total effect on employment following the boom in mineral prices. This of course, remains a descriptive result, but is telling in the sense that most of the effect on employment is driven by an increase in the employment of locals.

Which locals? Columns 2 and 3 in panel b.2 allow further analysis. Out of the 0.09 percentage points effect on the employment of locals is filled by low skilled locals: the estimated coefficient for this category is 0.06 percentage points. High skilled locals also benefit: the estimated coefficient is 0.03. To put this into perspective it is worth looking at the average changes in the migration rates. The average change in the migration rate of locals between 1993 and 2007 was -0.01 percentage points: the proportion of employed locals experienced a reduction in general. The estimated coefficient of 0.09 signals a large effect on the share of employed locals. By skill categories, the estimated effects are also important. The average change in the share of employed locals with high skill was 0.02 percentage points, which is not too small compared to the estimated effect for mining districts: 0.03. The effect on low skilled locals is the largest when compared to the average change. From 1993 to 2007 the proportion of low skilled locals employed fell by 0.04 percentage points, while mining districts following the boom experienced a positive effect, 0.06.

The effect on new comers can be dissected in a similar fashion. The average change in

the proportion of new comers employed from 1993 to 2007 was -0.01. The change for high skilled was 0.004 while for low skilled, -0.02, the estimated effects for the total proportion of employed new comers was 0.02, 0.01 for high skilled and 0.01 for low skilled (although this is not statistically significant). Results, therefore, indicate an important effect in the composition of the employment by migration category, with the proportion of locals benefiting more from the boom in mineral prices.

How does this result look between industries? Table 3.8 shows this in detail for the change in the log number of people employed in the district. The first three columns show a clear effect on the growth of locals employed in the mining sector in the district after the increase in the mineral prices. The coefficient estimated in column 1, 0.96, suggests that mining districts experienced a growth in the number of employed locals in the mining sector that was 29.8% (0.96x0.31) higher than in non-mining districts for every standard deviation increase in the index of mineral prices. Such effect is slightly higher for high skilled workers, where the estimated coefficient is 1.02. The coefficient for low skilled locals employed in the mining sector is 0.70. Interestingly, the mining boom also attracted high skilled individuals who were born in the district and returned from somewhere else. This result is shown in column 5, with a coefficient of 0.33. Something similar can be interpreted for high skilled *new locals* (individuals not born in the district but who were living there since five years ago) with an estimated coefficient of 0.54 (column 8).

Table 3.8: Change in Migration Composition by Industry - Log Levels

	N22			N21			N12			N11		
	Locals			Returned			New Locals			New Comers		
	N	H	L	N	H	L	N	H	L	N	H	L
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
a. Mining												
ΔP	0.96***	1.02***	0.70**	0.30	0.33***	0.14	0.57*	0.54***	0.44	0.55	0.29	0.50
	(0.34)	(0.22)	(0.35)	(0.18)	(0.08)	(0.23)	(0.30)	(0.17)	(0.31)	(0.41)	(0.37)	(0.38)
R^2	0.35	0.39	0.33	0.27	0.28	0.22	0.23	0.28	0.19	0.24	0.27	0.22
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
b. Tradable												
ΔP	0.13	-0.14	0.08	-0.22*	-0.12	-0.22*	-0.38	-0.11	-0.31	0.05	0.10	-0.07
	(0.22)	(0.30)	(0.19)	(0.12)	(0.11)	(0.13)	(0.26)	(0.19)	(0.28)	(0.22)	(0.15)	(0.21)
R^2	0.28	0.24	0.28	0.22	0.21	0.21	0.22	0.21	0.22	0.22	0.18	0.22
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
c. Non-Tradable												
ΔP	-0.07	0.17	-0.18	-0.18	-0.18	-0.08	-0.09	0.33	-0.14	0.09	0.18	0.16
	(0.22)	(0.28)	(0.19)	(0.17)	(0.17)	(0.20)	(0.19)	(0.24)	(0.16)	(0.26)	(0.28)	(0.27)
R^2	0.43	0.32	0.39	0.27	0.24	0.24	0.32	0.24	0.28	0.30	0.30	0.23
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
d. Agriculture												
ΔP	0.05	0.06	0.04	0.45**	0.20	0.47**	0.36*	0.53***	0.33	0.37	0.37	0.34
	(0.15)	(0.25)	(0.14)	(0.18)	(0.18)	(0.18)	(0.21)	(0.17)	(0.22)	(0.24)	(0.29)	(0.23)
R^2	0.47	0.29	0.47	0.30	0.16	0.29	0.30	0.25	0.29	0.31	0.18	0.31
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124

Notes: [1] Data source: population censuses of 1993 and 2007. [2] All regressions include as controls: altitude, historical coefficient of variation of rainfall, distance to Lima (not included in the sample), the log of district population in 1993 and province fixed effects. [3] The dependent variables are the current population composition (share) of the district according to migration categories, by industry as indicated in panels. *Locals* are those individuals born in the district that were living there five years ago. *Returned* are individuals born in the district who were not living there five years ago. *New Locals* are individuals not born in the district but were living there five years ago. *New Comers* are individuals who were not neither born or residents of the districts five years ago. For each category N holds for total, H for high skilled workers and L for low skilled workers. [4] Errors clustered at province level, and coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

There is no statistically significant result in the tradable or non-tradable sectors.

Panel (d) shows the results for the agriculture sector. There is a positive effect for individuals who returned and found a job in the agricultural sector, 0.45, which interestingly is focused in low skilled workers, 0.47. For every standard deviation increase in the index for mineral prices, the number of low skilled workers in the agricultural sector grew 14.6% (0.47x0.31) faster in mining districts. Interestingly, there is also an effect in the number of *new locals* who found a job in the agricultural sector: 0.36, which is focused in the group of high skilled workers, 0.53.

Table 3.9 re-evaluate these results in terms of employment shares. Recalling that the total effect on the employment rate was 0.13, the sum of all coefficients for the total estimates

(columns 1, 4, 7 and 10) should add up to this number[77]. In that regard, this table is useful to understand the contribution of every sector to the total effect on employment. Column 1 in the set of estimates for the mining sector indicates that out of the 0.13 effect on total employment, 0.03 is explained by an increase in the employment of locals working in the mining sector, and 0.06 percentage points are explained by locals working in the agriculture sector. More interesting, all the effect on locals who work in the agricultural sector come from the low skilled category (column 3, panel d), whereas the increase in mining employment is mostly explained by high skilled locals. The rest of the coefficients are statistically insignificant or very close to zero. The later, again is a result in itself for the tradable and non-tradable sector. This results confirm if the tradable sector did not benefit from the mining boom it was not harmed either.

[77] Due to rounding it does not.

Table 3.9: Change in Migration Composition by Industry - Shares

	N22 Locals			N21 Returned			N12 New Locals			N11 New Comers		
	N (1)	H (2)	L (3)	N (4)	H (5)	L (6)	N (7)	H (8)	L (9)	N (10)	H (11)	L (12)
a. Mining												
ΔP	0.03**	0.02**	0.01**	0.00**	0.00***	0.00*	-0.00	0.00***	-0.00	0.02	0.01	0.01*
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.23	0.23	0.22	0.20	0.27	0.14	0.17	0.15	0.21	0.16	0.21	0.12
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
b. Tradable												
ΔP	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.26	0.29	0.26	0.18	0.08	0.20	0.14	0.20	0.16	0.10	0.14	0.10
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
c. Non-Tradable												
ΔP	0.00	0.01	-0.01	-0.00	-0.00	-0.00	-0.01	0.00	-0.01**	0.00	0.00	0.00
	(0.02)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
R^2	0.29	0.44	0.23	0.20	0.18	0.20	0.28	0.45	0.19	0.30	0.31	0.27
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124
d. Agriculture												
ΔP	0.06***	0.00	0.06***	0.01*	0.00	0.01**	0.01	0.00***	0.01	-0.00	0.00*	-0.00
	(0.02)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
R^2	0.30	0.49	0.29	0.34	0.08	0.34	0.21	0.28	0.21	0.70	0.16	0.71
Observations	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043	1043
Clusters	124	124	124	124	124	124	124	124	124	124	124	124

Notes: [1] Data source: population censuses of 1993 and 2007. [2] All regressions include as controls: altitude, historical coefficient of variation of rainfall, distance to Lima (not included in the sample), the log of district population in 1993 and province fixed effects. [3] The dependent variables are the current population composition (share) of the district according to migration categories, by industry as indicated in panels. *Locals* are those individuals born in the district that were living there five years ago. *Returned* are individuals born in the district who were not living there five years ago. *New Locals* are individuals not born in the district but were living there five years ago. *New Comers* are individuals who were not neither born or residents of the districts five years ago. For each category N holds for total, H for high skilled workers and L for low skilled workers. [4] Errors clustered at province level, and coefficients that are statistically different from zero are denoted by the following system: *10%, **5% and ***1%

To summarize, results indicate that the mining boom exerted a positive influence over the employment share of the districts within 100km distance to any large scale mining project in 1993. Employment grew 13% faster, which can be read as 4% faster increase by every standard deviation increase (0.31) in the index of mineral prices at the local level. As a share, low skilled employment contributed with 7% while high skilled employment contributed with 5%. Unemployment declined by 13% and was focused on the group of low skilled workers. The growth in the employment rate was due to higher employment rates in the mining and agriculture sectors. No effect for tradable and non-tradable industries. The change in mining employment was mainly explained by higher high skilled employment while in the case of the agriculture industry, low skilled employment explained the increase.

In migration terms, most the increase in the employment share was explained by a higher share of employed locals. The majority of them, in turn, were low skilled, working in the agriculture sector. Higher high skilled employment in the mining sector was also filled by locals.

In the following section I discuss possible explanations for these results.

3.5 Discussion: are these results evidence of *Dutch Disease*?

Are these results together evidence of *Dutch Disease*? At this point it would be useful to recall the framework proposed by Corden and Neary (1982) to understand the existence of *Dutch Disease*. The results presented in this chapter support the hypothesis in favor of a *spending effect*.

In light of Corden and Neary (1982) this happens in the absence of a *resource movement effect*. The boom in the mining sector increases the marginal products of factors employed in the sector, which consequently draws resources out of the other sectors, which is followed by further adjustments in the economy. One of these adjustment takes place in the real exchange rate (prices of traded over non-traded goods).

Since the mining sector in reality draws very little resources from the rest of the economy (its labor absorption is very low while most of the capital used there is imported), the *resource movement effect* is negligible and the majority of its impact happens through the *spending effect*.

This is what the results show, the higher real income resulting from the boom leads to more spending on local services which consequently increase their prices and a real appreciation with more adjustments following. Notably, the agricultural sector, which is non-tradeable, benefits from the boom in the mining sector both in terms of employment and wages. Clearly this is evidence of the *spending effect* operating in the local economy.

Is this enough to conclude that local economies in Peru subjected to mining booms suffered a *Dutch Disease*? Certainly there are some of the elements in the results, mainly the increase in income, but the lack of evidence of de-industrialisation is also there. This lack of evidence of de-industrialisation however may be due to the lack of industrial activity in the region under analysis. Tradable employment in the region of analysis accounts for 9% of total tradable employment in the country. Even so, tradable employment in the area un-

der analysis accounts for 6% of total employment in the same area. Nationally, the tradable sector accounts for 10% of total employment. Therefore the tradable sector in the highlands of Peru is very small, which may make difficult to detect any result pointing towards de-industrialisation.

Together these results suggest that there are some elements of the *Dutch Disease* (mainly a *spending effect*) among the results and the lack of evidence of de-industrialisation cannot be taken as clear cut evidence of no harm to the industrial sector because the industrial employment in the region under analysis is very low.

3.6 Conclusions

This chapter explored the effects of large-scale mining activity on local labor markets. It proposed a simple spatial framework that accounts for heterogeneous industries and workers. In its simple setting, the model predicts a larger labor response of high skilled workers for a given elasticity of substitution between high skilled and low skilled labor. As a consequence, their wage response is smaller than the wage response of the low skilled workers.

Using census data, I find that the employment rate grew faster in districts close to large-scale mining activity. Both, high skilled and low skilled employment rates grew, and the low skilled unemployment rate decreased.

The population size grew similarly in both types of districts, but there was a re-shuffle in the proportion of high and low skilled workers. The first increased in detriment of the second.

In terms of industries, the effect on employment focused on the mining and agriculture sectors. The increase in agriculture employment is explained by an increase in employment for low skilled workers. While for mining the effect is explained by high skilled labor. This is evidence of a *spending effect* in the light of the *Dutch Disease* framework since most of the agricultural sells locally.

Using individual data I find that wages increased with the mining boom, which again is evidence of a *spending effect*. The evidence is suggestive of an heterogeneous effect on the skill and industry, however it is not statistically different. By industry, agricultural workers gained the wage increase. The gain experienced by the agricultural workers may be

explained by the large concentration of low skilled workers in this sector.

The chapter also explored the composition of population and employment according to the migration status of the individuals. Results indicate that most the effect on employment steam from higher employment for low skilled locals, mainly in agriculture.

As a concluding remark, this chapter sets the basic formulation to properly understand the geographical spillovers of industry booms in developing countries, however further extensions in the theoretical formulation should seek to endogenise the acquisition of education and the migration decision of individuals.

Final Remarks

This thesis collected three chapters on empirical topics associated to local labor markets, mineral booms, beliefs, conflict and uncertainty. All the analysis was conducted using Peruvian data and context. The degree of connection among the topics is high for the first two chapter, while the last chapter represents an analysis that dominates Peruvian economic debate.

The first chapter evaluated the effect of conflict over the formation of trust and identity. The chapter found that Peruvian individuals exposed to violent events during their impressionable years trust less government institutions, and feel less identified with their neighbors, while more identified with religious groups. The estimated effect is small and heterogeneous depending on the identity of the perpetrator. The effect on identification with groups of population is also heterogeneous by the indigenous origin of the individuals. People who own an agricultural plot embedded in a cooperative setting at the local level exhibit even smaller levels of identification with their locals while higher levels of identification with their ethnic group. In line with recent literature, these findings suggest that conflict has a small but persistent effect on the formation of trust and identity, which is a central feature to understand the interaction between culture and institutions, and ultimately to understand the persistent consequences of wars.

The second chapter analyzed the relationship between democratic beliefs and economic uncertainty. In particular, I explored whether uncertainty experienced during the impressionable years of the individuals is a key factor behind the formation of the democratic beliefs. Results showed that this type of uncertainty had no effect on the determination of democratic beliefs. Combining uncertainty with the exposure to authoritarian regimes did not change the result. This result is robust to different definition of rural individuals, the interaction of uncertainty and degree of experienced authoritarianism, and different formative periods. Current uncertainty, on the other hand, was unable to fully explain the formation of democratic beliefs.

The final chapter investigated the local labor effects of mining booms. Using two rounds of population census for 1043 districts in Peru I documented that large-scale mining activity had a positive effect on local employment over 14 years. The effect was differentiated by in-

dustry, skill and migration status. Employment grew by 4% faster by one standard deviation increase in the mineral prices. Both high and low skilled workers enjoyed similar employment increase, however only low skilled workers experienced a decline in unemployment. Using data from 10 annual household surveys I found that, consistent with a model of heterogeneous firms and labor, wages for low skilled workers in districts close to the mining activity was 5% higher by every standard deviation increase in the index of mineral prices. Additional evidence with the census data suggested that to a large extent locals working in the mining or the agricultural sector filled the new employment opportunities. Together these findings suggest that large-scale mining activity increases the demand for mining and agricultural local employment, and the wages in the local economy.

As a final remark, I would like to signal some potential ways to improve the chapters in order to depict possible routes in the research arena.

Both first and second chapters could benefit from a theoretical framework that guides better the empirical results.

In particular, the first chapter found a small effect of the experience of armed conflict over the degree of identification with people in the locality, which is enhanced by the indigenous origin of the individual. The chapter also outlined some of the possible channels to explain the association between conflict and trust and identity. Therefore, the natural following step would be to test them empirically. Data availability of those channels has not been available, but perhaps future studies could find statistical sources to test them.

From the first chapter it was also suggestive that there is an important effect depending on the individual communal behavior of the individuals. This is clear when the effect of conflict interacted with the individual measure of participation in communal arrangements carries the effect, rather the interaction of the district communal land share with the experience of conflict.

The latter could also be a consequence of the closed nature of the communal organizations in Peru. For instance, when half of the district's agricultural land in under the tenancy of the communal organizations, that district is considered communal in the empirical analysis. That measure, however, fails to provide an accurate degree of cooperative behavior. It is possible that the district is populated by numerous communities acting in their own inter-

est, certainly with cooperation within them, but not between them. Future research should investigate this in more detail.

The second chapter arrived to the conclusion that neither uncertainty nor a rainfall shock mattered for the shaping of democratic beliefs in Peru. Re-assessing the question with different statistical tools offers no different result. However, this does not necessarily mean that democratic beliefs cannot be explained somehow. After all, preferences for democracy are another type of political attitudes, which are hardly randomly determined.

Subsequent analysis should look for another source of explanation. Perhaps economic recessions also explain the preferences for democracy. This constitutes an straightforward extension of the study of Giuliano and Spilimbergo (2014). However, and perhaps more promising, the formation of democratic beliefs could be determined by the degree of inequality that the individuals are exposed to. After all, democracy is grounded in a sense of political equality, since every vote counts the same. Would individuals who feel their peers earn more have a weaker democratic preference? Of course, finding an exogenous source of variation in inequality is not straightforward. However, there have been attempts to understand the effect of inequality over subjective well-being that could be extended to this context. In concrete, the work of Luttmer (2005) is a relevant reference to guide this research path. There, the author finds that controlling for the individual's income, higher levels of neighbor income are associated with lower levels of self-reported happiness.

An interesting venue for future research in the last chapter consists in understanding the role of migration costs in the determination of the results. How does the current infrastructure network mitigates the local response to a mineral boom? A proper understanding of the theoretical and empirical implications could follow the study of Morten and Oliveira (2016).

Additionally, the theoretical framework could include the housing market in order to predict the effects of mineral booms over real wages and amenities.

Also, as in Alcott and Keniston (2015), in order to disentangle the connection among industries, it is necessary to include the goods market into the theoretical framework. It is important to understand how the increase in local revenues determines the profits in the local firms, depending on their destination market.

With all these extensions, this chapter could provide a very detailed explanation of the

consequences of a mineral boom, under the framework of local labor demand shocks.

References

ACEMOGLU, D., S. JOHNSON, J. ROBINSON, AND P. YARED (2008): "Income and Democracy," *American Economic Review*, 98, 808–842.

ACEMOGLU, D., S. NAIDU, P. RESTREPO, AND J. ROBINSON (2014): "Democracy Does Cause Growth," *National Bureau of Economic Research*, 20004.

ACEMOGLU, D. AND J. ROBINSON (2006): *Economic Origins of Dictatorship and Democracy*, New York: Cambridge University Press.

ADHVARYU, A. AND J. FENSKE (2014): "Conflict and the Formation of Political Beliefs in Africa," *Household in Conflict Network*, 164.

ALCOTT, H. AND D. KENISTON (2015): "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America," *National Bureau of Economic Research*.

ALEMÁN, J. AND D. YANG (2011): "A Duration Analysis of Democratic Transitions and Authoritarian Backslides," *Comparative Political Studies*, 44, 1123–1151.

ALGAN, Y. AND P. CAHUC (2010): "Inherited Trust and Growth," *American Economic Review*, 100, 2060–92.

ARAGÓN, F. AND J. RUD (2013): "Natural Resources and Local Communities: Evidence from a Peruvian Gold Mine," *American Economic Journal: Economic Policy*, 5, 1–25.

BARRON, M., E. MIGUEL, AND S. SATYANATH (2014): "Economic Shocks and Democratization in Africa," *Political Science Research and Methods*, 2, 33–47.

BARTIK, T. (1991): "Who Benefits from State and Local Economic Development Policies," *W.E. Upjohn Insitute For Employment Research*.

BAUER, M., C. BLATTMAN, J. CHYTILOVÁ, J. HENRICH, E. MIGUEL, AND T. MITTS (2016): "Can War Foster Cooperation?" *Journal of Economic Perspectives*, 30, 249–74.

BELLOWS, J. AND E. MIGUEL (2006): "War and Institutions: New Evidence from Sierra Leone," *American Economic Review*, 96, 394–399.

——— (2009): "War and local collective action in Sierra Leone," *Journal of Public Economics*, 93, 1144 – 1157.

BERTRAND, M. AND S. MULLAINATHAN (2001): "Do People Mean What They Say? Implications for Subjective Survey Data," *American Economic Review*, 91, 67–72.

BLACK, D., T. MCKINNISH, AND S. SANDERS (2005): "The Economic Impact of the Coal Boom and Bust," *The Economic Journal*, 115, 449–476.

BLANCHARD, O. AND L. KATZ (1992): "Regional Evolutions," *Brookings Papers on Economic Activities*, 1–75.

BLATTMAN, C. (2009): "From Violence to Voting: War and Political Participation in Uganda," *American Political Science Review*, 103, 231–247.

BRATTON, M. AND G.-B. E (2016): "Do Trustworthy Institutions Matter for Development? Corruption, Trust and Government Performance in Africa," *Afrobarometer*.

BRUCKNER, M. AND A. CICCONE (2011): "Rain and the Democratic Window of Opportunity," *Econometrica*, 79, 923–947.

CABALLERO, J. AND E. ALVAREZ (1981): *Aspectos Cuantitativos de la Reforma Agraria: 1969-1979*, Instituto de Estudios Peruanos.

CEPES (2005): *Informativo Legal Agrario*, 21, Centro Peruano de Estudios Sociales.

CHEN, Y. AND S. X. LI (2009): "Group Identity and Social Preferences," *American Economic Review*, 99, 431–57.

CORDEN, W. M. AND J. P. NEARY (1982): "Booming Sector and De-Industrialisation in a Small Open Economy," *The Economic Journal*, 92, 825–848.

CVR (2003): "Comisión de la Verdad y Reconciliación: Informe Final," .

DEATON, A. (1997): "The Analysis of Household Survey: A Mricroeconomic Approach to Development Policy," *Baltimore: John Hopkins University Press*, 1877–1979.

DEL CASTILLO, L. (1992): "Â¿Tienen Futuro las Comunidades Campesinas?" *Centro Peruano de Estudios Sociales - CEPES, Debate Agrario*, 14.

——— (1997): "Propiedad Rural, Titulación de Tierras y Propiedad Comunal," *Centro Peruano de Estudios Sociales - CEPES, Debate Agrario*, 26.

DELL, M. (2010): "The Persistent Effects of Peru's Mining *Mita*," *Econometrica*, 78, 1863–1903.

DELL, M., B. JONES, AND B. OLKEN (2012): "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, 4, 66–95.

DHAL, R. (1971): *Polyarchiy: Participation and Opposition*, New Haven, CT: Yale University Press.

DURANTE, R. (2010): "Risk, Cooperation and the Economic Origins of Social Trust: An Empirical Investigation," *Mimeo*.

DURANTE, R., E. ZHURAVSKAYA, AND I. GROSFELD (2015): "Middleman Minorities and Ethnic Conflict: Evidence from Pogroms," *Mimeo. Science-Po*.

EIFERT, B., E. MIGUEL, AND D. N. POSNER (2010): "Political Competition and Ethnic Identification in Africa," *American Journal of Political Science*, 54, 494–510.

FAFCHAMPS, M., M. KOELLE, AND F. SHILPI (2015): "Gold Mining and Proto-Urbanization: Recent Evidence from Ghana," *World Bank, Working Papers*.

GALLUP, J. AND J. SACHS (2001): "The Economic Burden of Malaria," *American Journal of Tropical Medicine and Hygiene*, 64.

GALOR, O. AND O. OZAK (2016): "The Agricultural Origins of Time Preference," *American Economic Review*, 106, 3064–3103.

GIULIANO, P. AND N. NUNN (2013): "The Transmission of Democracy: From the Village to the Nation-State," *American Economic Review: Papers & Proceedings*, 103, 86–92.

GIULIANO, P. AND A. SPILIMBERGO (2014): "Growing Up in a Recession," *Review of Economic Studies*, 81, 787–817.

GOUDIE, I. AND M. GOUDIE (2007): "Who captures the marks for the Petersen estimator?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170, 825–839.

GUARDADO, J. (2015): "Land Tenure, Price Shocks and Insurgency: Evidence from Peru and Colombia," *Mimeo. University of Chicago*.

GUERRIERI, V., D. HARTLEY, AND E. HURST (2013): "Endogenous gentrification and housing price dynamics," *Journal of Public Economics*, 100, 45 – 60.

HABER, S. (2012): "Rainfall and Democracy: Climate, Technology and the Evolution of Economic and Political Institutions," *Mimeo. Standord University.*

HSIANG, S., M. BURKE, AND E. MIGUEL (2013): "Quantifying the Influence of Climate on Human Conflict," *Science*, 341.

HUNTINGTON, S. (1991): *Thee Third Wave: Democratization in the Late Twentieth Century*, Norman, OK: University of Oklahoma Press.

INEI (1998): *Estructura y Dinámicas del Espacio Agropecuario*, Instituto Nacional de Estadisticas.

——— (2015): "http://censos.inei.gob.pe/censos1981/redatam/DOC/FichaTecnicaBaseDatos.pdf," *Instituto de Estadísticas e Informática*.

INTERRARMD (2013): "http://www.rmg.se/Products/RawMaterialsData.aspx," .

KELLY, T. AND G. MATOS (2013): "Historical statistics for mineral and material commodities in the United States (2013 version)," *U.S. Geological Survey Data Series*, 140.

KING, R., S. BIRD, A. OVERSTALL, G. HAY, AND S. HUTCHINSON (2013): "Injecting Drug Users in Scotland, 2006: Number, Demography and Opiate-related Death Rates," *Addiction Research and Theory*, 21, 235–246.

KLINE, P. AND E. MORETTI (2014): "People, Places and Public Policy: Some Simple Welfare Economics of Local Economic Development Program," *Annual Review of Economics*, 6, 629–662.

KROSNICK, J. AND D. ALWIN (1989): "Aging and Susceptibility to Attitude Change," *Journal of Personality and Social Spsychology*, 54, 416–425.

LEÓN, G. (2012): "Civil Conflict and Human Capital Accumulation: Long Term Consequences of Political Violence in Perú," *Journal of Human Resources*, 47, 991–1023.

LOAYZA, N., A. MIER Y TERAN, AND J. RIGOLINI (2013): "Poverty, Inequality and the Local Natural Resource Curse," *World Bank, Working Papers*.

LUTTMER, E. F. P. (2005): "Neighbors as Negatives: Relative Earnings and Well-Being," *The Quarterly Journal of Economics*, 120, 963–1002.

MANRIQUE-VALLIER, D., M. PRICE, AND A. GOHDES (2013): "Multiple Systems Estimation Techniques for Estimating Casualties in Armed Conflicts," in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. by T. Seybolt, J. Aronson, and B. Fischhoff, Oxford University Press, chap. 9.

MASELLA, P. (2013): "National identity and ethnic diversity," *Journal of Population Economics*, 26, 437–454.

MENDELSOHN, R. (2007): "What Causes Crop Failure?" *Climatic Change*, 81, 61–70.

MIGUEL, E. (2004): "Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania," *World Politics*, 56, 328–362.

MIGUEL, E. AND S. SATYANATH (2011): "Re-examining Economic Shocks and Civil Conflict," *American Economic Journal: Applied Economics*, 3, 228–232.

MORETTI, E. (2010): "Local Multipliers," *American Economic Review: Papers & Proceedings*, 100, 373–377.

——— (2011): "Local Labor Markets," *Handbook of Labor Economics*, 4, 1237–1313.

MORETTI, E. AND P. THULIN (2013): "Local Multipliers and Human Capital in the United States and Sweden," *Industrial and Corporate Change*, 22, 339–362.

MORTEN, M. AND J. OLIVEIRA (2016): "Paving the Way to Development: Costly Migration and Labor Market Integration," Mimeo, Stanford University.

NOTOWIDIGDO, M. (2013): "The Incidence of Local Labor Demand Shocks," *Northwestern University. Mimeo*.

OXFAM (2009): "Mining Conflicts in Peru: Condition Critical," Oxfam, Oxfam.

PARDO, M. (1996): *Estudios sobre la provincia de Jauja*, Colección Populibros regionales, Ediciones José María Arguedas.

PAXSON, C. (1992): "Using Weather Variability to Estimate Response of Savings to Transitory Income in Thailand," *American Economic Review*, 82, 15–33.

PUTNAM, R. (1993): *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press.

ROBACK, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90, 1257–1278.

SACHS, J. AND A. WARNER (1995): "Natural Resource Abundance and Economic Growth," *National Bureau of Economic Research*.

——— (1999): "The Big Push, Natural Resource Boom and Growth," *Journal of Development Economics*, 59, 43–76.

SOTELO, S. (2015): "Trade Frictions and Agricultural Productivity: Theory and Evidence from Peru," *University of Michigan. Mimeo*.

SVOLIK, M. (2008): "Authoritarian Reversals and Democratic Consolidation," *American Political Science Review*, 102, 153–167.

TOPEL, R. (1986): "Local Labor Markets," *Journal of Political Economy*, 94, 111–143.

A Appendix: CVR Data

As discussed by Manrique-Vallier et al. (2013), the methodology relies on four strong assumptions:

1. Closed system: the lists (every catch in the fish example) refer to a closed system: N must refer to the same population in every draw.

2. Homogeneity: for each list (or draw), every individual (or fish) must have the same probability of being included, or captured.

3. Independence: the probability that an individual is included in list A is independent from the probability that the same individual is included in list B, and vice versa.

4. Random selection: each list is a random draw of the total population, N.

In the application of the methodology to the estimation of casualties, the assumption 1 is almost irrelevant, while assumption 4 is not possible, which requires some robustness checks. Assumption 2 is less worrisome, however it depends on the effort exerted in the collection of observations for every list. If the agents collection information of every list are independent, assumption 3 should hold.

In the context of the Peruvian case, the CVR defined 6 lists, or data sources, with information of the 24,692 reported victims:

1. CVR itself: information on the number of victims, based on 16,886 testimonies.

2. National Coordinator of Human Rights: for the period 1983-2000, this government institution collected information about victims reported to the institution across the country.

3. Center for Agrarian Development: specific of the number of victims in the district of Chungi, Ayacucho. The information was collected by the municipality of Chungi.

4. Commission for Human Rights: for the period 1982-1996, this government institution collected information about victims reported to the institution across the country. However, it has a particular focus in the regions of Ayacucho, Apurimac and Huancavelica.

5. Population's Advocate: this institution collects information of missing people, as well as executions. The information used covered the period 1983 to 1996.

6. Red Cross: list with information of missing people collected by the Red Cross.

The details of the estimation, as well as the treatment of the assumptions that led to the estimate of 69,280 can be found in the Appendix 2 of the Final Report elaborated by the CVR. At this point it is worth mentioning that the statistical procedure involve the estimation of models that predicted the probability of being included in all the lists[78] or some of them, by geographical strata and identity of the perpetrator. The statistical models calculated directly the probability of being a victim of the army, while the estimates for the number of victims attributed to PCP-SL, and the other groups was estimated as a residual from the total number of victims estimated and the victims estimated as casualties from the army.

B Trust and Identity Questions

Both questions are consistently available for the period 2007-2012. Here I report the translation of the questions

B.1 Trust

Table B.1: Question: Currently, do you trust institutions like...?

Institution	A lot	Enough	Little	Nothing	Do not know	National Average	% Do not Know
The National Police of Peru	1	2	3	4	5	0.1993	0.1163
The Armed Forces	1	2	3	4	5	0.3261	0.1574
The Judicial Power	1	2	3	4	5	0.1717	0.1620
The Parliament	1	2	3	4	5	0.1282	0.1832
Political Parties	1	2	3	4	5	0.0675	0.1498
The Catholic Church	1	2	3	4	5	0.5790	0.0870

I have re-coded categories 1 and 2 as 1 and categories 3 and 4 as 0, while excluded those individuals who responded Do not know. In the sample I include only individuals with all valid responses. The total number of individuals included in the analysis for the period 2007-2012 is 104,635. National average excludes "do not know" answers.

[78]The final number of lists was reduced to 3: CVR, Population's Advocate, and the rest. This helped the statistical procedure.

B.2 Identity

Table B.2: Question: With which group (or community) do you feel more identified with?

Group	Code	Average
Your region, province, district or ward	1	0.6128
Your ethnic group or race	2	0.1783
Your farmers or indigenous community	3	
Your religious group or faith	4	0.1942
Other group	5	0.0147

For the 104,635 with information on trust, I also derived information on the reference group. The question asked directly for identification with any of the groups listed. I collapsed categories 2 and 3 since they can be used interchangeably and also to gain statistical representation.

C Survey questions on democracy

Table C.1: Opinion on the Importance of Democracy

Sample	**2002-2004**	**2004-2006**	**2007-2012**
Question	*In your opinion, the importance that democracy has for a government is?*	*In your opinion, the importance that democracy has for a government is?*	*In your opinion, democracy in Peru is?*
Answers	**A lot = 1**	**A lot = 1**	**Very Important = 1**
	More or less = 0	Enough = 0	Important = 0
	Little = 0	Little = 0	Little important = 0
	Nothing = 0	Nothing = 0	Not important = 0
	Do not know = .	Do not know = .	Do not know = .
National	0.456	0.426	0.196
Rural	0.348	0.328	0.131

Notes: [1] Data source: ENAHO 2002-2012. [2] Unweighted averages.

Table C.2: Type of Preferred Government

Sample	**2002-2006**	**2007-2012**
Question	*In your opinion, the political government that is more convenient to rule the country is?*	*With which of the following statements do you agree the most?*
Answers	Authoritarian = 0	**A democratic government = 1 is always preferable**
	Military = 0	Sometimes an authoritarian is = 0 preferable than a democratic one
	A democratic government = 1	It is the same to me = 0
	The experts, not the gvt = 0	Do not know = .
	Do not know = .	
National	0.605	0.565
Rural	0.550	0.491

Notes: [1] Data source: ENAHO 2002-2012. [2] Unweighted averages.

Table C.3: Despite its Imperfections, Democracy is the Best Option for a Government

Sample	**2002-2004**	**2004-2006**
Question	*Do you share the following opinion: democracy can cause trouble, but is better than any other form of government?*	*Do you share the following opinion: democracy has defects, but is better than any other form of government?*
Answers	**Much = 1**	**Much = 1**
	More or less = 1	**Enough = 1**
	Little = 0	Little = 0
	Nothing = 0	Nothing = 0
	Do not know = .	Do not know = .
National	0.655	0.459
Rural	0.534	0.343

Notes: [1] Data source: ENAHO 2002-2012. [2] Unweighted averages.

Table C.4: Democracy Works

Sample	**2002-2004**	**2004-2012**
Question	*In the country, does democracy work?*	*In the country, does democracy work?*
Answers	**A lot = 1**	**Very well = 1**
	More or less = 1	**Well = 1**
	Little = 0	Bad = 0
	Nothing = 0	Very bad = 0
	Do not know = .	Do not know = .
National	0.553	0.354
Rural	0.549	0.403

Notes: [1] Data source: ENAHO 2002-2012. [2] Unweighted averages.